Bridging the Valley of Death in Biomedicine with Translational Research: Assessing the Impact of National Institutes of Health's Clinical and Translational Science Award

A Dissertation
Presented to
The Academic Faculty

by

Yeon Hak Kim

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Public Policy

Georgia Institute of Technology

August 2019

Bridging the Valley of Death in Biomedicine with Translational Research: Assessing the Impact of National Institutes of Health's Clinical and Translational Science Award

Approved by:

Dr. Aaron Levine, Advisor
School of Public Policy
Georgia Institute of Technology

Dr. Eric Nehl
Department of Behavioral Sciences and Health Education
Emory University

Dr. Alan Porter
School of Public Policy
Georgia Institute of Technology

Dr. Weihua An
Department of Sociology
Emory University

Dr. John Walsh
School of Public Policy
Georgia Institute of Technology

Date Approved: May 15, 2019

ACKNOWLEDGEMENTS

I want to express my deepest appreciation to my advisor, Dr. Levine. Throughout my Ph.D. journey, you have always guided me to the right path and supported me with kind words. Thanks to your help, I learned how to do good research and I will never forget our weekly meetings in your office. I would also like to thank my committee members, Dr. Porter, Dr. Walsh, Dr. Nehl, and Dr. An, who always taught me great lessons whenever I reach out to them and provided invaluable comments. It was my honor to have you in my committee and I would not have completed this journey in good shape without your support. I also want to thank Dr. Llewellyn for her support that helped me get thorough understanding of my dissertation topic.

I am very grateful to the people I met at Georgia Tech. Thank you, Dr. Marco, Dr. Massetti and Dr. Klein, who taught me a lot of lessons during our research meetings that helped me improve my research skills. I always felt that I was being cared whenever I was having meetings with you. I am also grateful to Ms. Charnigo for her kind and friendly administrative support. I would also want to thank my friends, Karim Farhat, Evan Mistur, Simone Rosenblum, Quintin Kreth, Usayd Casewit, Anmol Soni, Yeong Jae Kim, Elie Sung, Seok Beom Kwon and Seokkyun Woo, who made my life at Georgia Tech enjoyable.

Finally, I would like to thank my family. To my lovely wife Junghee, thank you so much for all your support and love. To my precious son, Joonie, thank you for coming into my life and being healthy. Also, I would like to express my appreciation to my parents and my sister, who always trusted me and encouraged me throughout my whole life. I would also like to thank my parents-in-law, who gave me constant emotional support. I really love you all!

CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES viii

LIST OF FIGURES x

SUMMARY xii

CHAPTER 1. INTRODUCTION 1

1.1. Research Motivation 1

1.2. Research goals and objectives 2

1.3. Overview of the chapters 4

CHAPTER 2. POLICY BACKGROUND 6

2.1. Overview of the biomedical research pipeline 6

2.1.1. Complex and multi-stage process 6

2.1.2. Various stakeholders with heterogenous roles 9

2.2. The challenge: decreasing productivity and the 'valley of death' in biomedicine 15

2.2.1. Eroom's law 15

2.2.2. Valley of death in biomedical research 18

2.3. Translational research: A new solution for "valley of death" in biomedical research 20

2.3.1. The concept of translational research 20

2.3.2. Support for translational research in the United States 22

2.3.3. Support for translational research in other countries 28

CHAPTER 3. CHANGE OF COLLABORATION AMONG RESEARCH UNIVERSITIES INFLUENCED BY SUPPORT FOR TRANSLATIONAL RESEARCH 31

3.1. Introduction 31

3.2. Literature review 32

3.2.1. Factors affecting research collaboration and their relevance with the CTSA award .. 32

3.2.2. The impact of the CTSA award on research collaboration network 35

3.3. Theoretical background and hypotheses 37

3.4. Data and methodology 39

3.4.1. Empirical setting 39

3.4.2. Measures and variables 49

3.4.3. Models 54

3.5. Results 56

3.5.1. Descriptive analysis 56

3.5.2. Network regression 65

3.6. Conclusion 70

CHAPTER 4. A MEASURE ON TRANSLATIONAL FEATURE OF PUBLICATIONS 72

4.1. Introduction 72

4.2. Background 73

4.3. A new measure: the TS score 76

4.4. Verification tests on the measure 78

4.4.1. Reliability tests 78

4.4.2. Validity tests 82

4.5. Conclusion 90

CHAPTER 5. IMPACT OF NIH'S CTSA AWARD ON TRANSLATIONAL RESEARCH PERFORMANCE OF INSTITUTIONS 92

5.1. Introduction 92

5.2. Theoretical background, previous studies and hypotheses 93

5.2.1. The direct relationship between the CTSA award and translational outcome 93

5.2.2. The indirect relationship between the CTSA award and translational outcome 97

5.3. Data and methodology 102

5.3.1. Empirical setting 102

5.3.2. Data and variables 103

5.3.3. Method 110

5.4. Results 113

5.4.1. Institution matching 113

5.4.2. Descriptive analysis 117

5.4.3. Estimation results 123

5.5. Conclusion 131

CHAPTER 6. DISCUSSION 134

6.1. Summary of findings 134

6.2. Implications 136

6.3. Limitation and suggestion for further studies 139

APPENDIX A. List of CTSA hubs and their partnering institutions 142

APPENDIX B. Networks measures with x-axis as normalized years 144

APPENDIX C. Goodness-of-fit assessment of binary ERGMs 148

APPENDIX D. Classification of basic science, clinical science, multidisciplinary science, and other science based on NSF's classification of fields of study 150

APPENDIX E. More examples the change in TS score by year for some publications 154

APPENDIX F. Institution matching based on propensity score of taking advantage of the CTSA award 158

APPENDIX G. Data summary and correlation between variables having all Carnegie R1 universities in the sample 160

APPENDIX H. More DID regression results for test on Hypothesis 1 161

APPENDIX I. Counts and portion of CTSA supported papers by institution-year 164

APPENDIX J. Model estimation results for testing mediation effects 168

REFERENCES 169

LIST OF TABLES

Table 1. Institutions classified by leadership in the CTSA program (as of 2015) 42

Table 2. Mean values of R&D related attributes for R1 and R2 universities in 2014 43

Table 3. Concepts of centrality measures 51

Table 4. List of control variables used in the analytic model 53

Table 5. ERGM estimations with binary dependent variable between 2009 and 2015 66

Table 6. ERGM estimations with weighted dependent variable between 2009 and 2015 68

Table 7. Estimation result of the TERGM for 2009-2015 69

Table 8. Correlation between TS score of the starting point and TS score after four or more years 81

Table 9. Correlation between TS scores calculated with different methods 84

Table 10. TS score comparison between publications resulting and not resulting in patents 87

Table 11. Summary of TS scores of Emory University's non-clinical publications that result in patents 88

Table 12. TS score of non-clinical science publications that result in drugs 90

Table 13. The expected direction of the impact of factors on the change of scientist's behavior towards translational research 96

Table 14. Descriptions of the variables used in the analysis 109

Table 15. Required conditions for medation effect to be present 112

Table 16. CTSA selection rules and potential factor affecting the selection 115

Table 17. Group size based on CTSA starting year 116

Table 18. Data summary with comparison of CTSA-evers and CTSA-nevers 118

Table 19. Mean translational publication counts before and after CTSA for all institutions 120

Table 20. Mean translational publication counts before and after CTSA based on institution matching with 0.5 caliper option 123

Table 21. Estimation of main difference-in-difference regression models (dropping unmatched institutions) 125

Table 22. Estimation of difference-in-difference regression models excluding publications with direct CTSA support 128

Table 23. Result of mediation tests taking inter-organizational collaboration as mediator..... 129

Table 24. Result of mediation tests taking multidisciplinarity as mediator..... 130

LIST OF FIGURES

Figure 1. The traditional process of biomedical product development 8

Figure 2. Annual FDA approvals and total R&D Expenditures in 2010 dollars 16

Figure 3.Average FDA approvals per 2010 US billion dollars in R&D..... 16

Figure 4. Illustration of the "Valley of Death" 19

Figure 5. Major events of NIH's support for translational research 24

Figure 6. Number of organizations with the history of receiving the CTSA award 25

Figure 7. Map of CTSA-funded institutions 26

Figure 8. Change of the sizes of the treatment group and the control group by year 45

Figure 9. Flow chart of making collaboration network between Carnegie R1 universities 48

Figure 10. The trend of density and centralization of the whole network 57

Figure 11. Centralization and density for the network between CTSA-ever institutions..... 59

Figure 12. Centralization and density for the network between CTSA-never institutions..... 59

Figure 13. Collaboration frequency with CTSA-evers 60

Figure 14. Collaboration frequency with CTSA-nevers 61

Figure 15. Collaboration portion with CTSA-evers 61

Figure 16. Collaboration portion with CTSA-nevers 62

Figure 17. Degree centrality trend of institutions in CTSA-ever group and CTSA-never group. 63

Figure 18. Closeness centrality trend of institutions in CTSA-ever group and CTSA-never group 63

Figure 19. Betweenness centrality trend of institutions in CTSA-ever group and CTSA-never group 64

Figure 20. Eigen-vector centrality trend of institutions in CTSA-ever group and CTSA-never group 64
Figure 21. Classifying publications into the ones with and without translational feature based on journal discipline of forward citing publications 77
Figure 22. Change in TS score by year for some non-clinical publications 80
Figure 23. Classifying publications into the ones with and without translational feature based on first author's affiliation 83
Figure 24. Distribution of TS scores calculated using different approaches 85
Figure 25. The depiction of Hypothesis 2 98
Figure 26. The depiction of Hypothesis 3 102
Figure 27. The empirical setting of my analysis in chapter 5 103
Figure 28. Trend of translational publication counts by calendar year 119
Figure 29. Trend of translational publication counts based on institution-matching without caliper option 120
Figure 30. Trend of translational publication counts based on institution-matching with caliper 0.05 option 122
Figure 31. Average TS score by publication group 126
Figure 32. Change of the portion of CTSA award acknowledging publications by year 126
Figure 33. Trend of multidisciplinary score of institutions 130

SUMMARY

Despite large investment in biomedical research by government, foundations and private organizations around the world, we are not experiencing an increase in the new medicine reaching the market. Many studies point out that this productivity decline in biomedicine is mainly due to the difficulty in translating basic science into clinical setting. Translational research emerged as a key research policy tool to address this problem over the last decade. Translational research aims to bridge the gap between basic science and clinical science to accelerate the process of moving research innovation into clinical use. In the United States, the National Institutes of Health (NIH) took the lead in supporting translational research by developing the Clinical and Translational Science Award (CTSA) in 2006. In this dissertation, the author examined the impact of NIH's effort on supporting translational research focusing on two topics, which are collaboration network structure and production of translational publications. Regarding the collaboration landscape, the change of social network analysis measures showed that the CTSA award had an impact in changing the biomedical research landscape into denser and less centralized form. The result of the network regression models showed that receiving CTSA award led individual institutions to collaborate more with other institutions. For the test on the production of translational publications, which is the second topic of interest, a unique measure using the composition of forward citation of publications is introduced. The results from difference-in-difference regression and mediation tests showed that the CTSA award leads to the increase of publications and this relationship is mediated by inter-collaboration feature of institutions after the CTSA program is well stabilized. The author expects that the study will provide insight into the effects of translational research initiatives and have implications on the government policy regarding biomedical research more broadly.

CHAPTER 1. INTRODUCTION

1.1. Research Motivation

Translational research emerged as a key research policy issue over the last decade. This was a response to a series of studies concluding that remarkable advances in basic biomedical research were not leading to a significant increase in the development of new medicines (Butler, 2008; Moran, 2007; Zerhouni, 2005; Woolf, 2008). Translational research aims to fill the gap between basic science and clinical science to accelerate the process of moving research innovation into clinical use (Fishburn, 2013; The President's Council of Advisors on Science and Technology, 2012). Translational research initiatives use several strategies to reach these goals, including providing tools that promote collaboration among institutions and researchers, building research infrastructure with multiple purposes, supporting pilot programs, training students and faculty members, and providing regulatory support (CTSA Principal Investigators, 2012; Lander & Atkinson-Grosjean, 2011).

Within the United States, the National Institutes of Health (NIH) took the lead in supporting translational research by announcing its Roadmap Interdisciplinary Research Initiative in 2005, developing the Clinical and Translational Science Award (CTSA) in 2006 and establishing the National Center for Advancing Translational Sciences (NCATS) in 2011 (Blümel, 2017; Gittelman, 2016; Obeid, Johnson, Stallings, & Eichmann, 2014). NIH awarded its first twelve CTSA awards in 2006 and now supports around sixty institutions nationwide providing money for the instrumentation, infrastructure, training, pilot projects and regulatory support (Clinical and Translational Science Awards Consortium, n.d.-a). Individual CTSA

receiving institutions are funded usually through 5-year agreements and their annual budget ranges from $4 million to $23 million[1] (Leshner, Terry, Schultz, & Liverman, 2013). NCATS, which had a budget of over $500 million in 2017, houses the CTSA program and many other programs that facilitate translational research (Liu, Chen, Sinoway, & Berg, 2013).

Given the scale of NIH's support for translational research and interest in and support for translational research efforts around the world, it is important to understand if these research policies are meeting their goals. By systematically assessing the effects of NIH's CTSA program, I aim to increase understanding of the impact of translational research policies and identify strategies to improve the design and implementation of biomedical research funding programs. Also, I aim to improve the evaluation of programs designed to foster translational research by proposing and completing preliminary assessments of the reliability and validity of a novel measure to assess the translational nature of research articles based on their forward citation profiles.

1.2. Research goals and objectives

In this dissertation, I aim to examine the impact of NIH's primary effort to support translational research, the development and implementation of the CTSA Award. My aim is to test several hypotheses on how NIH's support for translational research changed the collaborative network structure in biomedical research and how, if at all, production of biomedical research articles – a key output of the research enterprise – changed due to this transformation. Here, my goal is to find out, in particular, whether NIH's support for

[1] For instance, Emory University received $104.7 million of sponsored research funding in Fiscal Year 2016 and 6% ($6.2 million) of them were related to the CTSA program.

translational research was associated with the production of research articles that were used in a translational manner. My specific research questions are as follows:

- Motivating questions

1) Has the government's support for translational research through the CTSA program changed the overall biomedical research collaborative landscape among research universities in the U.S.?
2) Has the NIH's increasing support for translational research led to more production of research outcomes that translate basic science findings into clinical setting? If so, what are the mediating factors in the relationship?

Thus far, several studies have looked at the impact of support on translational research, mainly with basic descriptive analysis focusing on individual institutions (e.g., Liu et al., 2013; Knapke, Haynes, Kuhnell, & Tsevat, 2015). However, there has been less focus on conducting comparative assessment among different organizations using archival data. In that regard, the objectives of my study are as follows:

- Specific objectives

1) Quantitatively analyze if NIH's support for translational research changed the features of biomedical research collaboration landscape using social network (SNA) analysis measures and models.
2) Analyze the archival data related to publications and use appropriate multivariate regression techniques to assess the impact of NIH's policy on the characteristics of research outcomes and institutions' research performance.

3) Contribute to the literature on the effectiveness of government intervention in biomedical research collaboration and performance.

1.3. Overview of the chapters

The remaining parts of the dissertation will be structured as follows. In chapter 2, I will provide an overview of the policy background relevant to this dissertation. First, I will explain general aspects of biomedical research then explain how these features lead to the need for collaboration in biomedical research. In particular, the presence of the "valley of death" in biomedical research will be explained with the reasonings why it is present. Then, I'll discuss a new approach to biomedical collaboration, which is translational research. First, I will explain the concept and briefly explain the background that led to the emergence of this type of research. Then, I will provide information on the efforts of the U.S government regarding the support of translational research. Although the focus of this dissertation is translational research policy in the United States, for comparative purposes, I will briefly introduce similar policies in other countries in Europe and Asia.

In chapter 3, I will investigate how the collaborative landscape among research intensive universities changed due to NIH's support for translational research. Following a discussion of key literature on research collaboration, focusing on studies that examine the factors influencing the level of collaboration and how external forces induce research network changes, and relevant theoretical background, I provide an explanation of my empirical setting. The multi-step process of developing the dataset for this analysis is detailed and explanation of the variables and models used for the assessment of the CTSA program is provided. Then, the descriptive analysis of

network measures, and the results of network regression models will be provided. The chapter concludes with a discussion of the results of my analysis, along with interpretation of these results.

The focus of chapter 4 is the development and preliminary assessment of a novel measure on the translational feature of publications. First, background on why a new measure on translational feature of publications is necessary will be presented. Then, a detailed explanation of the new measure – the TS score – that measures the translational feature of individual articles or institutions will be provided. Finally, results on the tests on the reliability and the validity of the new measure will be described.

The goal of chapter 5 is to extend the analysis from chapter 3 and to assess the relationship between research performance and support for translational research. Drawing on the review of related theories and previous studies, I hypothesize that the NIH's support for translational research changed scientist's behavior to produce research outcomes that can be considered translational publications. Using the TS score to determine translational publications (described in chapter 4), I test this hypothesis by econometrics methodologies like difference-in-difference methodology in various settings. Furthermore, tests on the mediation effects of two major factors that NIH emphasizes for the successful operation of the CTSA program, inter-organizational collaboration and multidisciplinary research, are tested. Then, the result with descriptive analysis of key variables and model estimation results are presented. In closing of this chapter, the discussion on the results will be provided.

Finally, chapter 6 will provide a summary of the findings, implication of the study, and limitation with suggestion on further studies.

CHAPTER 2. POLICY BACKGROUND

2.1. Overview of the biomedical research pipeline

2.1.1. *Complex and multi-stage process*

In biomedical research, there are multiple stages to go through to translate a discovery from the laboratory to a product with commercial value. The first step, in the typical development of new drug[2], is the discovery and development process, which usually occurs in the laboratory (U.S. Food and Drug Administration, 2018). In this stage, researchers conduct tests to find promising compounds that may stop or reverse the effect of a particular disease. If a promising compound is identified, researchers conduct experiments to characterize the compound and determine, for example, an effective delivery mechanism, an appropriate dosage, and its interaction with other drugs. The second stage is preclinical research (U.S. Food and Drug Administration, 2018). In this stage, researchers seek information on dosing and toxicity levels of potential drugs by conducting in vitro and in vivo tests including animal tests. When the tests suggest that the potential drug has a reasonable chance of successful development and passed basic tests on drug's safety, an Investigational New Drug (IND) application, which includes information such as preclinical study data, manufacturing information, clinical protocol, is submitted to the FDA (Niosi, 2011; Fishburn, 2013). FDA staff review the submitted information and make a go - hold decision, which can take as long as six months (Niosi, 2011). If the FDA

[2] Here, I mean typical development of new drugs as drugs developed from classic small molecule drugs, which make up more than 90% of commercial drugs (Bayer AG, n.d.). The processes of developing drugs using substances that consist of complex mixtures (e.g., biological drugs, cell therapies) are more challenging (Crommelin et al., 2003) and the features (e.g., study size, timeline and percentage of success) will differ with the contents provided in here.

review team approves the IND submission, the process moves to the next step: clinical research. Clinical research is composed of phase 1, phase 2 and phase 3. Phase 1 studies typically recruit 20 to 100 healthy volunteers to test the safety and dosage of the potential drug (U.S. Food and Drug Administration, 2018). It takes several months or more, depending on the nature of the project, to pass through this phase and around 70% of potential drugs go to phase 2. In phase 2, the study participants increase to several hundred people and tests are conducted to test the efficacy and side effects of the potential drug. Several months to two years are needed to pass through this stage and approximately 33% moves to phase 3 (U.S. Food and Drug Administration, 2018). Phase 3, which takes as long as four years and as many as few thousand participants, tests the efficacy and monitor the adverse reaction attributed to the potential drug. Around 25 to 30 percent of tested drugs pass this phase (U.S. Food and Drug Administration, 2018). When all phases of clinical trials are complete and if the potential drug is proved to have efficacy and safety, the drug developer can submit a New Drug Application (NDA) to FDA for the approval that allows them to sell the drug to individual customers. After a thorough review, FDA decides whether to approve or disapproves NDA. If NDA is approved, the drug can go to the market and reach the marketplace. In some cases, post-market drug safety monitoring, which is sometimes expressed as phase 4 clinical trial, take place. The prevailing model of biomedical product development is provided in Figure 1.

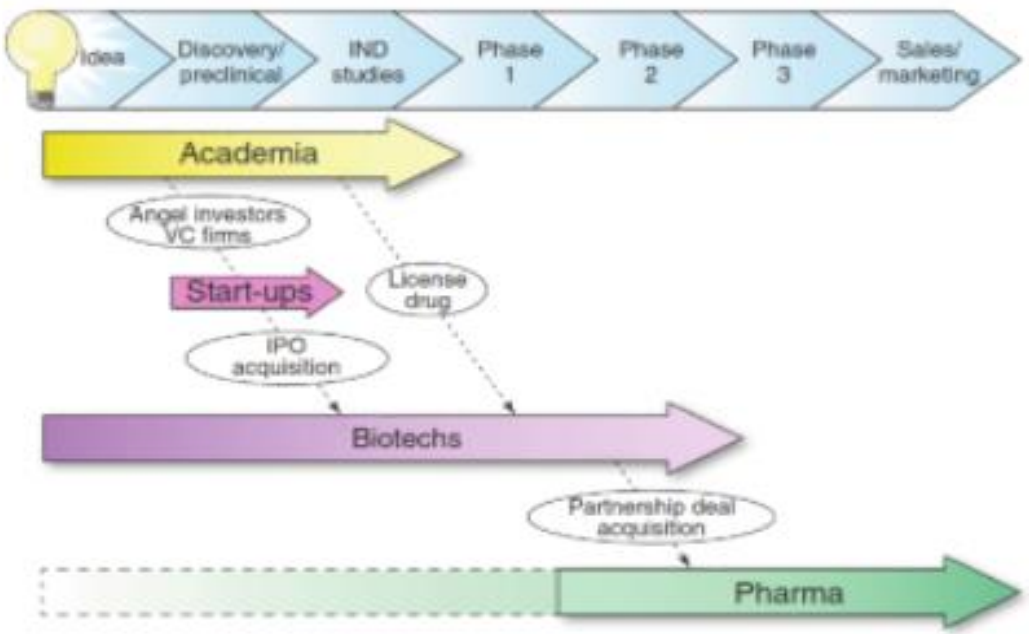


Figure 1. The traditional process of biomedical product development

Source: Fishburn (2013, p. 489)

Though abovementioned process is probably the most widely taken, there are new approaches that aim to accelerate the process of drug development (Knoepfler, 2015). For instance, FDA adopted paths to expedite the approval process by using surrogate endpoint as the evidence of safety and efficacy (e.g., accelerated approval mechanism), or putting more resources for the review (e.g., priority review mechanism) (Knoepfler, 2015). FDA also has some experimental regulatory approach such as Fast Track, which aims to speed up the approval process of breakthrough therapies that have been proven to have substantial enhancement over current therapies with preliminary experiment data (Knoepfler, 2015). Additionally, in some rare cases, FDA authorizes the use of unapproved drugs for the patients with terminal illnesses, which is referred as 'compassionate use' (Knoepfler, 2015; Zettler & Greely, 2014). The patient's physician and drug company both have to agree on the use of the unapproved drug when submitting proposal to FDA (Knoepfler, 2015). FDA reviews the process in a case-by-case basis

and gives approvals in case the proposal shows enough sign of safety and if there are no other comparable treatment options for the patient (Zettler & Greely, 2014).

2.1.2. *Various stakeholders with heterogenous roles*

The field of biomedicine is a representative field to which the concept of the National Innovation System (NIS) can be applied (Niosi, 2011). As there are multiple stages in biomedical product development, there are various stakeholders involved. Three major players traditionally mentioned in the NIS system, which are government, university, and private firms, all participate actively in the field (Niosi, 2011). In addition, there is a unique player in biomedical research arena that makes it distinct from other fields of science, which is the hospital (Hicks & Katz, 1996; Lander & Atkinson-Grosjean, 2011). Though the division of institutions in biomedical research arena is continuously changing, having a good understanding on the traditional roles of each player would help us understand the complex translation mechanism of biomedical products (Lynskey, 2006). Hence, I provide brief descriptions of players in each sector as below.

Government. In the case of the U.S., federal government agencies include regulatory agencies like FDA and funding agencies like NIH. As described in detail in previous section, FDA is the agency that approves the market entry of a drug. It is generally considered to have the most rigorous drug approval process in the world, allowing it to function as the global standard in many cases (Nam, 2015). Though the high level of requirements from FDA would ensure the safety and efficacy of drugs, there is a call from the public for timely delivery of new biomedical products that can bring favorable balance of benefits and risk (Califf & Ostroff, 2015). This call reflects a variety of factors, including patient frustration with the pace of drug development as

well as the development of alternative regulatory models, such as Japan's adoption of expedited conditional approval system in regenerative medical therapies[3], which allows certain products to gain market access with only limited evidence of efficacy. In the similar vein, FDA is adopting new regulation tools to accelerate the process of translating discoveries into approvals such as accelerated approval system, priority review mechanism and introduction of compassionate use, which are described in the prior section (Califf & Ostroff, 2015; Knoepfler, 2015). This trend led to the enactment of the 21st Century Cures Act in 2016, which allocates substantial support to FDA for rapid drug approval process. This act allows active use of biomarkers, observational data and surrogate measure to streamline the drug approval process (Gabay, 2017).

NIH oversees discovery research and clinical research (Fishburn, 2013). It is the largest source of government funding for biomedical research in the U.S. and the global biomedical market (Fishburn, 2013). In 2018, NIH had a total budget of $37.4 billion (National Institutes of Health, 2017c). The NIH's funding is awarded to researchers and institutions across all stages of the biomedical research pipeline from idea generation to clinical research (National Institutes of Health, 2017c). NIH also has its own laboratories that house around 6,000 scientists. In addition to traditional stages of research process, it also supports projects that aim to translate findings from basic research into clinical research, which is the focus of this dissertation and will be explained in detail later.

In the U.S., state governments also play an important role in the medical research ecosystem (Feldman, Lanahan, & Lendel, 2014; Nam, 2015). Individual states have adopted a wide variety of programs to support scientific research, including biomedical research, within their borders (Feldman et al., 2014). These include programs to fund biomedical research to

[3] In 2014, the Pharmaceutical and Medical Devices (PMD) Act created an expedited conditional regime and Act of the Safely of Regenerative Medicine (ASRM) allowed cell processing to be outsourced (Japan External Trade Organizations, 2017).

advance promising science and return economic benefits (Feldman et al., 2014; Karmali, Jones, & Levine, 2010). For instance, states like California, Connecticut, Illinois, Maryland, New Jersey and New York have their own funding programs that support stem cell research (Karmali et al., 2010, Alberta, Cheng, Jackson, Pjecha, & Levine, 2015). In some cases, states make their own laws or regulations regarding the use of biomedical products. For example, the right-to-try bill took into effect first in Colorado in 2014, which allows the use of drugs that passed phase 1 clinical trials, but without final approval from FDA, for the patients in severe condition[4] (Zettler & Greely, 2014). We could expect that these conditions that only apply to a small subset of states would have an impact on the rate and speed of translation process of biomedical research.

Universities. Academic institutions, which are mainly research universities in the U.S context, take the role of conducting various kinds of studies including research on basic science. In the U.S., research universities started to be established in the 1880s, and there was a massive increase in the size from 1990s to 2000s, and now the number of organizations reached about 350 in 2010 (Fishburn, 2013).

One of the major features regarding the structure of biomedical research network is that academic institution is the center of the network (Balconi, Breschi, & Lissoni, 2004; Lander, 2013; Owen-Smith & Powell, 2003, Schummer 2004). Lander (2013) found that universities are the key players in the network that connects hospitals, government organizations and private firms together. In the similar vein, a study by Schummer (2004) showed that academic institutions dominate the biomedical research arena and they mediate the interaction between governmental research institutions and the industry. Also, Owen-Smith and Powell (2003) showed that academic institutions are the central player in the inter-organizational knowledge

[4] This bill expanded to 38 states as of 2017 and was signed into federal law in May 2018 (Brown, Ortiz, & Dubé, 2018)

transfer network in life science. All these studies show that universities are the central locus of innovation in biomedical research and suggest that it would be worth examining academic institutions carefully in order to understand the general aspects of biomedical research landscape.

There are many reasons that universities are the center of the biomedical research network. One of the reasons is that universities produce the most advanced technology in biomedicine. Together with publicly funded research institutes, universities have been producing the technologies in the frontiers (Lynskey, 2006). These technologies led to the production of valuable biomedical products and had made universities to be recognized as the leaders of the regional innovation system (Lynskey, 2006). The other reason is related to the emergence of the new mission of the university, in which the universities take the expanded role of community engagement beyond their traditional roles of teaching and research (Lander, 2013). Universities are getting more engaged in economic development activities (Lander, 2013). This led academic institutions to interact more with external partners, including private sector and governmental organizations (Etzkowitz et al., 2008). By having more links with industry, universities could learn more about how the industrial product development process works, and what topics are important to that product development process (Fries, Glave, & Radick, 2008). This will narrow the gap between two sectors and this would lead to more coherent and effective collaboration between the two (McElroy, Jones, & Barrault, 2017).

Private firms. In the private sector, pharmaceutical companies take major roles in the drug development process. The total R&D spending of pharmaceutical companies (PhRMA member companies) were $71 billion (Statista, 2019). The pharmaceutical industry is also a large employer, employing an estimated 650,000 globally (Fishburn, 2013). Beyond marketing and sale of commercial drugs, pharmaceutical firms also conduct research. One of the ways they get

engaged in this process is by partnering with academic institutions (Roy & Chaguturu, 2014). This kind of partnership benefits both parties as they can exchange resources and make translation of ideas easier (Munos, 2014).

Other major participants in the private sector are biotech firms, which have been around since the 1970s (Fishburn, 2013). They are defined as the companies using biological organisms, systems or processes for drug development itself or providing a platform for new drug development (Huggett, Hodgson, & Lähteenmäki, 2011). In many cases, biotech firms take major roles in basic research and early stages of clinical development (phase 1 and phase 2) (Fishburn, 2013). They take advantage of translational research efforts that universities are conducting to enhance their chance of success in the early stage of drug discovery (McElroy et al., 2017). Biotech firms raise their funds through partnership deals with pharmaceutical companies, public market offerings and private investment (Fishburn, 2013; Huggett et al., 2011; Lynskey, 2006).

Other important players in the private sector of biomedical research arena are the investors, commonly referred to as angel investors and venture capitalists. Angel investors typically provide seed money to help the biotech firms move beyond pre-seed or seed stage, which typically requires $25,000 to $2 million (Eynott & Fages, 2014). Venture capital firms invest at various stages in the product development process, from the early stage of research to later stages much closer to commercialization. However, private investors are increasingly moving away from early-stage investing and don't wish to take the risk of investing in technologies with uncertain commercial potential. Data shows that only 4% of venture capital funds went into early-stage companies in 2013 (Eynott & Fages, 2014). In this regard, it is

becoming more difficult for biotech companies to execute successful commercialization of the technologies they possess.

Hospitals. In the clinical pathway of innovation, biomedical technologies must be tested to become a product with commercial value, and hospitals are where the research outcomes of the academic institutions can be applied and tested (Lander, 2013). However, the role of the hospital is often overlooked when discussing biomedical research system and the network among hospitals, universities and their resources are sometimes considered as "hidden research system" (Lander & Atkinson-Grosjean, 2011). Even in the non-commercial sector of biomedical R&D where the novelty of the research outcomes is important, the role of hospitals can become more significant (Lander, 2013). As Dosi (1982) noted, non-commercial organizations like academic hospitals may be open to sharing their findings and be more enthusiastic in collaborating with others if they think they can benefit from the interaction (Lander, 2013). In this regard, considering hospitals as a major player in biomedical research network would overcome the approach of the framework that only focuses on the commercial actors (Lander, 2013) and helps us better understand the translation process in biomedical research.

We can see that there are numerous stakeholders involved in the biomedical research pipeline. However, substantial differences between organizations in terms of their goals and missions exist and, hence, some obstacles hinder the smooth progress of the drug development (Eynott & Fages, 2014). Therefore, the success of biomedical research will depend on how well each stakeholder deals with this complex environment of diverse players and how we improve the collaboration among them.

2.2. The challenge: decreasing productivity and the 'valley of death' in biomedicine

2.2.1. *Eroom's law*

The number of new drugs approved by FDA per billion dollars spent has halved approximately every nine years since the 1950s (Scannell, Blanckley, Boldon, & Warrington, 2012). This decreasing trend of productivity in drug development is represented by the term 'Eroom's law,' which is the word that arranges the letters of 'Moore's law' backward (Scannell et al., 2012). Figure 2 illustrates the decreasing productivity of drug development. We can see that since the 1970s the rate of new drug approval is in a decreasing trend or in a constant state (shown by bars) despite the exponential growth in total R&D expenditure in drug development by the industry (shown by line) (The President's Council of Advisors on Science and Technology, 2012). The number of the FDA approved drugs, which is the sum of New Molecular Entity (NME) and the New Biologic Entity (NBE) approvals, tend to fluctuate without any sign of increase. In addition to the non-increasing trend of drug approvals, the number of new drug application to FDA is also in a decreasing trend (Borenstein, 2011). If we see the change of average number of FDA approvals per 2010 billion dollars spend in pharmaceutical R&D (Figure 3), we can clearly note that there is a decreasing trend of R&D productivity from mid-1970s.

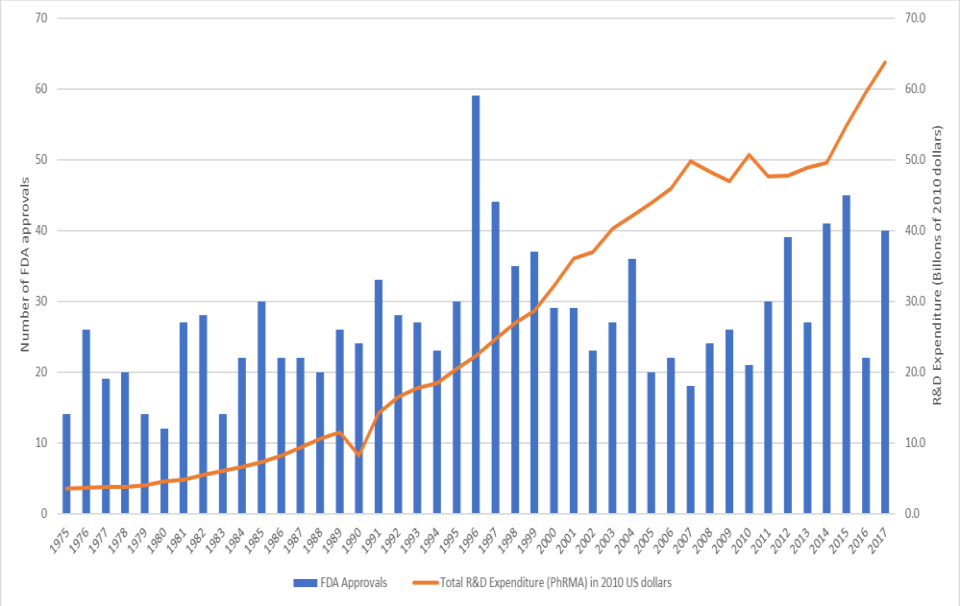


Figure 2. Annual FDA approvals and total R&D Expenditures in 2010 dollars

Note: Calculated with data from FDA and PhRMA

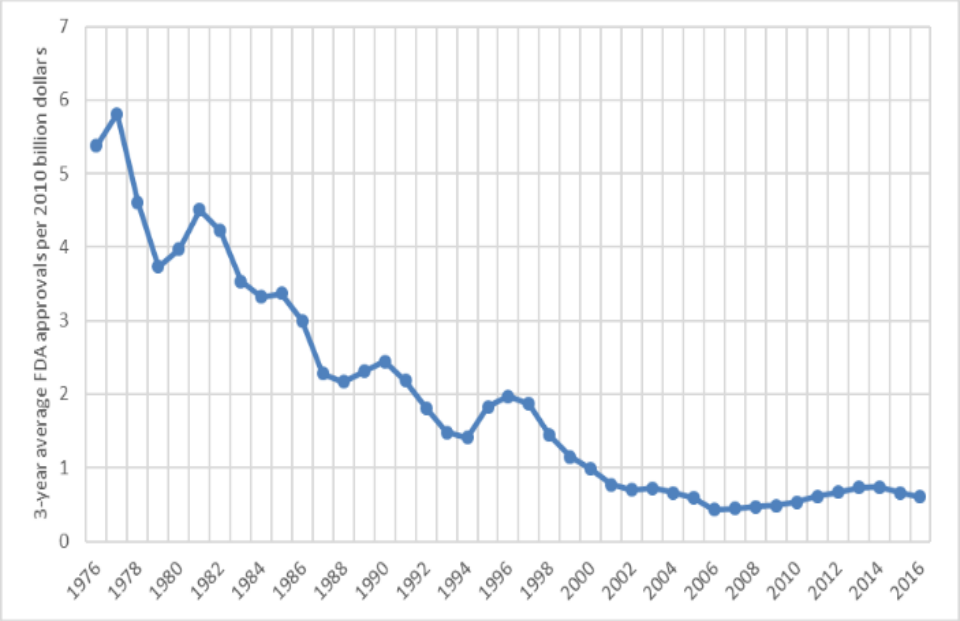


Figure 3.Average FDA approvals per 2010 US billion dollars in R&D

Note: Calculated with data from FDA and PhRMA

The decrease of biomedical R&D productivity is mainly attributed to two main factors, which are correlated with each other. The first factor is that lengthy time is needed for the translation of the discoveries into the product at the bedside. Morris, Wooding and Grant (2011) stated that the average time between the invention of discovery in the basic side of the research and the creation of something that is clinically useful is 17 years. By looking at the time needed from basic science discovery to the publication of first highly cited human research article, Contopoulos-Ioannidis, Alexiou, Gouvias, & Ioannidis (2008) showed that the median value of this time lag is 24 years. For some special cases, it may even take around 30 years to go through this process (Wilhelm et al., 2016). Also, in recent years, new type of therapies that are different from the traditional treatments using small molecule and biologics are emerging (Dodson & Levine, 2015). The introduction of this new concept of therapies, like cell therapies, make regulatory processes difficult to predict and could lead to a longer development process than before (Dodson & Levine, 2015).

The second related factor is the increasing cost of drug development. Kinch and Hoyer (2015) stated that the cost of discovering a new drug in the U.S. has been increasing and is now $2.6 billion. Some of the many factors contributing to the rise of the drug development costs are unmet medical needs becoming more complex, call for larger group of patients needed for clinical trials, stricter rules and more extensive testing, and emergence of rare diseases (Kinch & Hoyer, 2015; Nam, 2015; Wild, Huwe, & Lessl, 2013). In addition, novel therapies, such as cell therapies, could encounter high cost due to lack of pre-existing tools, reliable production line and distribution logistics (Dodson & Levine, 2015). The increasing cost puts investors at risk and may encourage risk-averse behavior when making investment decisions. Therefore, private funding is moving more towards the safer target (e.g., later stage proof-of-concept), which

subsequently led to less progress in the early stage of R&D that is essential to the productivity growth (Eynott & Fages, 2014; Kinch & Hoyer, 2015).

2.2.2. *Valley of death in biomedical research*

Usually, the early-stage of biomedical research pipeline and later-stage of the development do not suffer from extreme investment shortage (Steinmetz & Spack, 2009). The entities in public sector (e.g., federal government, foundations and state programs) invest heavily in early-stage biomedical research and universities and national laboratories actively participate in this stage of R&D (i.e., discovery and basic science). The funding for the later stage of the development comes from the private sector (e.g., big pharmaceutical companies, public-private partnerships and venture capital firms) and private firms voluntarily take part in the later stage of the process (i.e., product development and production). On the other hand, the cash flow between these two parts falls dramatically, leading to longer time between basic research and late-stage research, which is termed the 'valley of death.' (Ciensiski, 2015; Eynott & Fages, 2014; Steinmetz & Spack, 2009). The existence of the 'valley of death' is attributed to the fact that the chance of getting a significant return on research at this stage is not high and it requires a lot of investment (Ciensinski, 2015; Eynott & Fages, 2014).

The origin of the concept of the valley of death is from studies on innovation in general (Butler, 2008; Ciensinski, 2015). It illustrates the gap between basic science and product development (Figure 4). The step between basic science and product development is applied research, which translates the outcomes from basic research into a commercially valuable product (Ciensinski, 2015; Eynott & Fages, 2014). In biomedical research pipeline, the stage of

the valley of death is where the preclinical development, early clinical trials and some parts of early commercialization take place, which is essential for the FDA approval of biomedical products (Steinmetz & Spack, 2009). Due to monetary shortage in these stages, researchers get limited access to resources that are essential in technology commercialization procedures (e.g., scientific and business expertise, facilities, grant money) (Steinmetz & Spack, 2009). This in turn leads to a large portion of outcomes from basic research being not translated into actual products with commercial value.

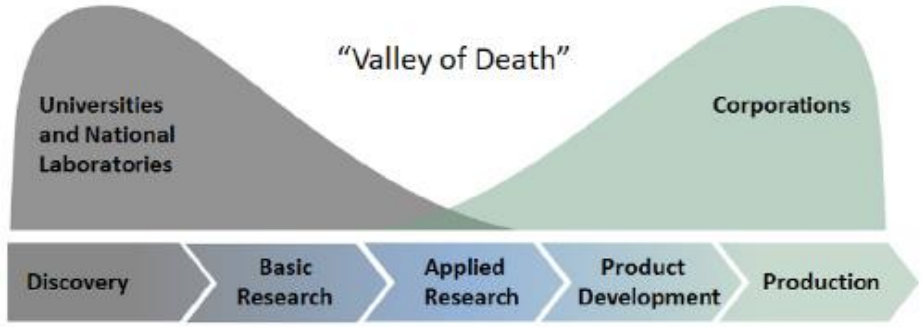


Figure 4. Illustration of the "Valley of Death"

Source: Ciensinski (2015)

As explained, the concept "valley of death" is mainly used to describe the monetary shortage. However, the concept is not explained by a single factor (Butler, 2008). It is a comprehensive concept that involves various factors that slow down the process of passing through the commercialization chasm. For instance, Butler (2008) argues that one of the reasons that led to the creation of the "valley of death" is the difference between the incentive system of basic scientists and that of clinical scientists. He claimed that for the basic scientists, their assessment largely depends on the quantity and the quality of the academic journal papers in their field, whereas physicians are assessed largely by other factors like the number of patients

they treated. Hence, basic scientists prefer to stay in their comfort zone of conducting research that can result in basic science publications and do not wish to participate in works related to clinical science (Blümel, 2017; Butler, 2008).

Unfortunately, what is evident is that the problem of the "valley of death" is not likely to be solved without the intervention from the public sector. As stated in the earlier sections, private investors are investing increasingly less on biomedical technologies in early commercialization stage to take risk and invest more on the technologies that are proven to have high potential of being commercialized (Eynott & Fages, 2014; Kinch & Hoyer, 2015; Moran, 2007). In this regard, to solve the problem, the research community needed a new approach with players in the public sector taking the leading roles.

2.3. Translational research: A new solution for "valley of death" in biomedical research

2.3.1. *The concept of translational research*

To tackle the problem of low productivity in biomedical research, the research community sought a new solution. As mentioned in previous sections, a series of studies pointed out that remarkable advance in basic biomedical research has not led to significant increase of new medicine, and many pointed out the "valley of death" as the core reason for this phenomenon. As a result, the concept of translational research attracted the interest of the biomedical research community as it had the purpose to fill and bridge the gap between basic science and clinical science (The President's Council of Advisors on Science and Technology, 2012). Having started to gain popularity in the 1990s (Lander & Atkinson-Grosjean, 2011), translational research has been promoted over the past couple of decades as a potential tool to

speed up the slow process of moving research innovations into clinical interventions and acceptance[5] (Fishburn, 2013; Stevenson et al., 2013).

One other reason that led to the emergence of translational research is to answer the questions from the society at large. As various societal issues in the society need science for their solutions, there has been continuous request from the society that science should get involved (Blümel, 2017). Translational research can be seen as a response to this call for societal impact of science and it aims to produce knowledge that benefits the patients and ultimately the society as a whole (Blümel, 2017). In this regard, we can see that the goal of translational research seems clear. Its aim is creating new research outcomes that are closely related to patient needs and translating discoveries in the laboratory into new clinical therapies (Fishburn, 2013; Surkis et al., 2016). The concept and definition of translational research are provided by institutes like NIH, Institute of Medicine (IOM), Translational Research Working Group of the National Cancer Institute (NCI). They all share the similar aspects and goals of increasing efficiency and reducing the time of transfer of fundamental scientific research into clinical research by bridging the "gap between the bench to bedside" (Rubio et al., 2010; Han, Williams, & Zuckerman, 2018).

The core feature of translational research is that it needs to bring people together for the success (Lander & Atkinson-Grosjean, 2011). As translational science span a wide range of boundaries of science, breaking down academic barriers, particularly between disciplines, is essential in the process (Lander & Atkinson-Grosjean, 2011). Many scholars claim that scientists in universities and hospitals should collaborate with the common goal of bridging basic science

[5] One specific example of translational research is a study by Bhat et al. (1997) that found binding of fusion genes (BCR-ABL) and certain type of protein (c-CBL) only occurs when phosphate (PO_4^{3-}) is added to acid (amino tyrosine) on a protein. Their finding was applied in treating genetic abnormality in chromosome 22 of leukemia cancer cell and led to the invention of Tasiga® (Sampat and Pincus, 2015). Another example of translational research is a study by Garg and Hassid (1990). They found out that proliferation of cell lines developed from disaggregated mouse embryos (BALB/c 3T3) is more active when muscle smoother is not present (CGMP-independent mechanism). Their finding was applied in solving respiratory failure problem and this led to the invention of INOmax® (Sampat and Pincus, 2015).

and clinical practice (Lander & Atkinson-Grosjean, 2011). By doing so, the chance of generating discoveries that can benefit both basic science and clinical science will increase (Lander & Atkinson-Grosjean, 2011).

Scholars and practitioners sometimes characterize translational research as consisting of multiple stages. One approach divides translational research into two steps. This approach claims that translational research is usually composed of two bidirectional stages (Drolet & Lorenzi, 2011; Han et al., 2018; Rubio et al., 2010; Kim, 2013; U.S. National Library of Medicine, 2017). The first stage, usually denoted as T1, involves transferring results from early-stage basic research into clinical research and the second stage, usually denoted as T2, involves spreading the use of research outcomes from clinical studies in actual practice and settling in local communities (Rubio et al., 2010). The second approach takes translational research as a four-stage process (Weber, 2013). These four stages are diagnosis or treatment (T1), evidence-based research (T2), clinical practice (T3) and verification for actual practice (T4) (Lander & Atkinson-Grosjean, 2011). The third approach, presented in the study by Surkis et al. (2016) and originally from reports from NCATS and Institute of Medicine, assumes that there are five steps, which are basic biomedical research (T0), translation to humans (T1), translation to patients (T2), translation to practice (T3) and translation to communities (T4). In this dissertation, I will focus on the stages related to the research activities (e.g., T1 in two-stage approach, T1 and T2 in four-stage approach and T0, T1 and T2 in five-stage approach).

2.3.2. *Support for translational research in the United States*

A brief history

The private sector, represented by biotech firms, has tried to narrow the translational gap for a long time but it turned out they failed to accomplish their goal completely (Gittelman, 2016). Since the early 2000s, NIH realized the problem and started taking action in supporting this 'gap filling' research (Gittelman, 2016). In the U.S., NIH Director Elias Zerhouni announced that the NIH would emphasize translational research as a part of its roadmap in 2003 (Han et al., 2018). In this roadmap, Zerhouni (2003) claimed that translational research could be a solution to solve uncertainty problems in clinical research. In 2005, NIH unveiled its Roadmap Interdisciplinary Research Initiatives and announced that the organization would put more effort into promoting translational research (Huerta et al., 2005; Lander & Atkinson-Grosjean, 2011). After the announcement of related roadmap and initiative in 2003 and 2005 respectively, NIH created an award dedicated to supporting translational research named Clinical and Translational Research Award (CTSA) in 2006 (Blümel, 2017). The NIH designated twelve institutions per year on average in the early period as the recipient of CTSA (Obeid et al., 2014). In 2012, six years after the creation of CTSA, an independent organization specialized in managing translational research named the National Center for Advancing Translational Science (NCATS) was established[6]. The budget of NCATS has increased from $545 million in 2013 (1.8% of total NIH budget) to $742 million in 2018 (2.0% of total NIH budget)[7], which is one evidence that NIH is in the process of increasing its support for translational research. Figure 5 shows major events regarding the NIH's increasing support for translational research.

[6] NCATS was created with the dissolution of the National Center for Research Resources (NCRR) that managed CTSA programs. It was founded to deal with the topics like 1) Bridging Interventional Development Gaps (BrlDGs), 2) Clinical and Translational Science Awards (CTSA), 3) Cures Acceleration Network, 4) FDA-NIH Regulatory Science, 5) Office of Rare Disease Research, 6) Therapeutics for Rare and Neglected Diseases (Rockey, 2012).

[7] Calculated from data in NIH website (https://www.nih.gov/about-nih/what-we-do/nih-almanac/appropriations-section-2)

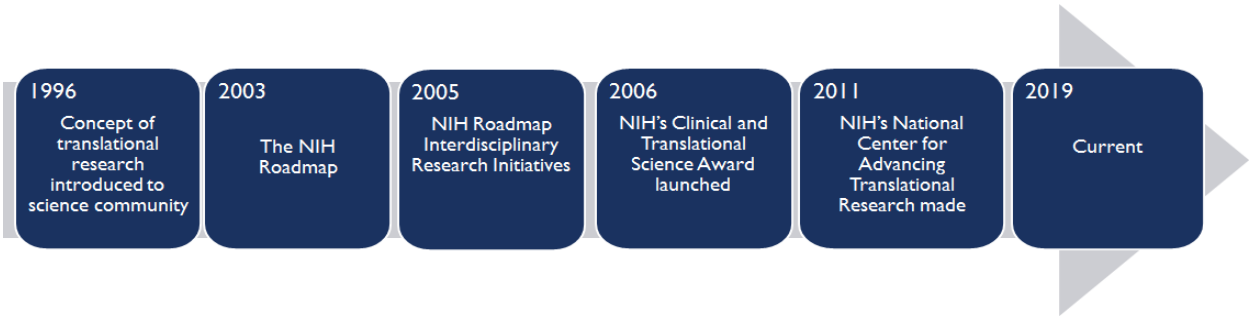


Figure 5. Major events of NIH's support for translational research

The CTSA program

The CTSA program is the heart of NIH's effort to support translational research. The institutions that receive the CTSA award establish Clinical and Translational Science Centers (CTSC)[8] and become the member of national NIH CTSA consortium (Butler, 2008), which shared $500 million in the fiscal year 2016 (Surkis et al., 2016). The consortium has the shared goal to enhance the efficiency and the quality of the translational research (Clinical and Translational Science Awards Consortium, n.d.-a). The size of the consortium increased from 12 in 2012 to 63 in 2015 (Figure 6 and Appendix A). CTSCs' services are not limited to technical support. They also 1) provide infrastructure with multiple uses, 2) provide tools for collaboration, 3) support pilot programs, 4) provide education and training, 5) provide administrative support (Leshner et al., 2013; CTSA Principal Investigators, 2012). To get a better understanding of the program, some details of each element is provided below. First two factors are related to both clinical and translational research, while other factors are more closely related to translational research.

[8] CTSCs can be regarded as an enhanced version of the General Clinical Research Centers (GCRCs) that started their operation in 1959 in universities and medical centers, which mainly focused on conducting clinical trials (Leshner et al., 2013). The budget of 78 GCRC sites in 2005 was $288 million, which is around 1.0% of total NIH budget that year (Leshner et al., 2013).

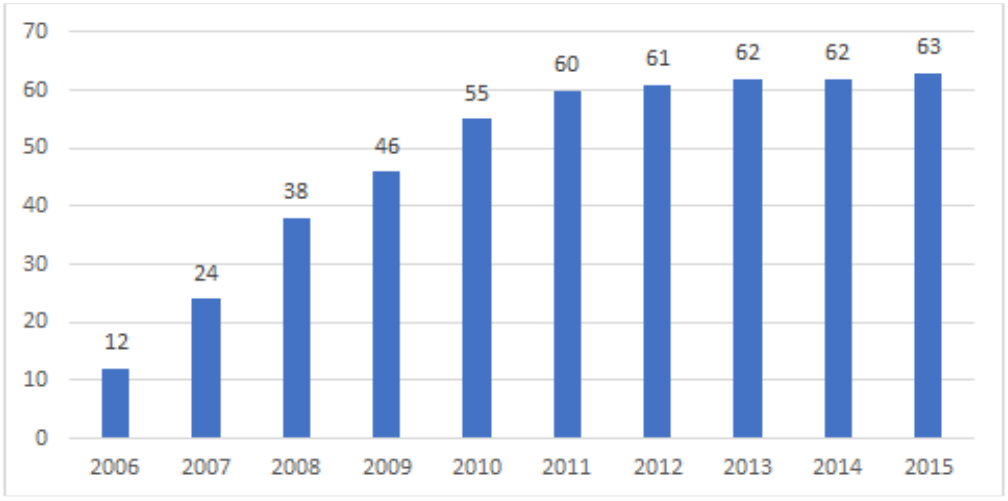


Figure 6. Number of organizations with the history of receiving the CTSA award

Infrastructure with multiple uses. Broad and reusable facilities that can be used in multiple disease domains are provided. It has been reported that 5,866 NIH grants in 2010 used facilities made with CTSA grant money (Leshner et al., 2013). Centers, such as Clinical Research Centers (CRC), which provide facilities for laboratory, nursing care and more, are established in multiple sites with the funds in this category (Leshner et al., 2013). CRCs provide resources to researchers that plan to conduct clinical studies that span across all stages of drug development. CRCs also facilitate drug repurposing studies (CTSA Principal Investigators, 2012)

Administrative support. CTSA receiving institutions also provide administrative support to streamline research (e.g., to reduce the time for Institutional Review Board (IRB) protocol review). For example, electronic sharing resource (e.g., IRBshare) is provided to the researchers to reduce the length of time needed for IRB approval award (CTSA Principal Investigators, 2012). Also, as a part of administrative support, support on contract negotiation and support on the recruitment of clinical trial participants are also provided (CTSA Principal Investigators, 2012).

Provide tools for collaboration. Figure 7 shows the map of where CTSA receiving institutions are located. We can see that the centers are located all over the nation with the focus on the regions that traditionally stand out in the field of biomedicine (e.g., California, Massachusetts, and Maryland). The CTSA program provides tools that help collaboration among these organizations and beyond. For instance, a web-based system that helps to locate experts in various disciplines (e.g., Profiles Research Network software, VIVO) are established (CTSA Principal Investigators, 2012). Also, systems for the data collection and information sharing are being operated (e.g., CTSA Intellectual Property Portal, Pharmaceutical Assets Portal) (CTSA Principal Investigators, 2012).

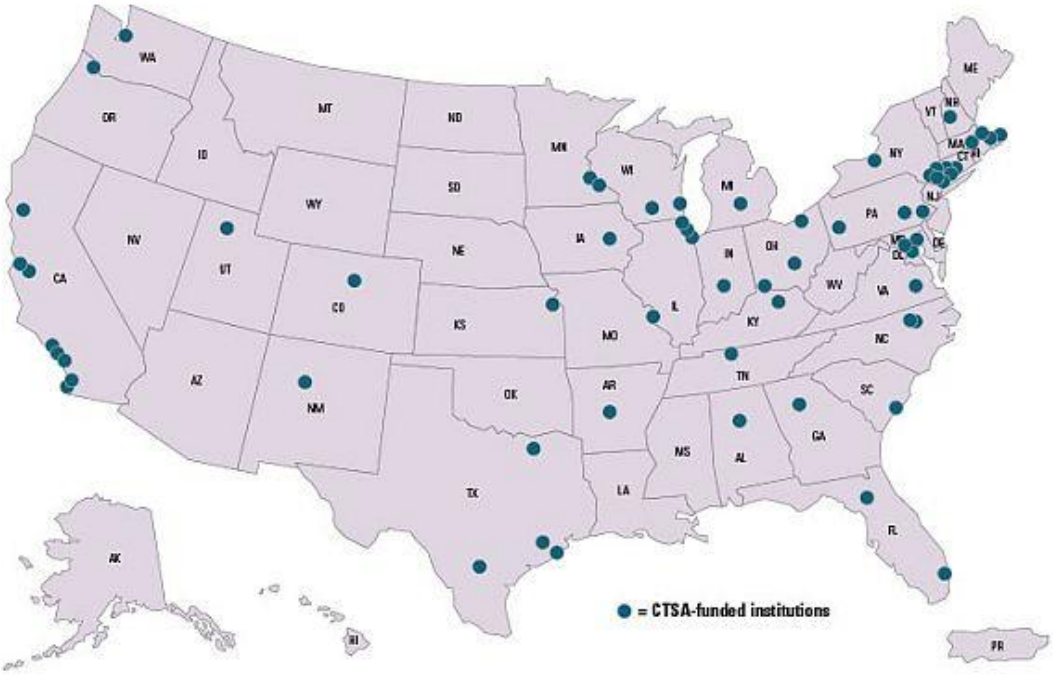


Figure 7. Map of CTSA-funded institutions

Source: Clinical and Translational Science Awards Consortium (n.d.-a)

In addition to these technical tools, an initiative called the Trial Innovation Network (TIN) was formed within the CTSA program in 2016 to facilitate more collaborations between institutions within CTSA recipients (Clinical and Translational Science Awards Consortium, n.d.-b). TIN is composed of three key organizational partners, which are 1) the CTSA Program Hubs, 2) the Trial Innovation Center (TIC), and 3) the Recruitment Innovation Center (RIC). There are three TICs nationwide, based at Johns Hopkins University, the University of Utah and the Duke University - Vanderbilt University Medical Center Partnership[9] (Slagle, 2016).

Support pilot programs. CTSAs typically allocate a portion of their funding to support small-scale pilot projects (National Institutes of Health, n.d.). This is to help researchers generate preliminary data for grant proposals and develop innovative methods and technologies (CTSA Principal Investigators, 2012; Leshner et al., 2013). The process of allocating pilot grants is designed to be flexible and simple so that it can provide opportunities to junior researchers and help them develop their research abilities (CTSA Principal Investigators, 2012). In 2011, around 2,000 pilot studies were conducted in various CTSA receiving institutions (CTSA Principal Investigators, 2012).

Education and training. One of the CTSA award's aims is to launch new educational initiatives to strengthen clinical research capabilities and improve the career of clinical researchers (Weber, 2013; Luke et al., 2015). To do so, CTSA provides training programs for acting clinical scientists with supplementary KL2 award (CTSA Principal Investigators, 2012; Knapke et al., 2015; Surkis et al., 2016). The CTSA program also supports master's degree

[9] Duke University – Vanderbilt University Medical Center partnership is a unique case in the sense that it is the only case with multiple leading institutes (Slagle, 2016). The partnership is between Duke Clinical Research Institute (DCRI) and Vanderbilt University Medical Center (VUMC), and the grant is planned to last for seven years with grant size of $26.5 million (Slagle, 2016).

programs in clinical and translational research with TL1 award (CTSA Principal Investigators, 2012).

Each CTSA receiving institution sets priorities for their funding based on their specific situation[10] but, all in all, the CTSA program aims to change the working environment of the whole institution. Activities associated with CTSA awards can support essentially all researchers and projects affiliated with the award-receiving institution. CTSA awards do not concentrate on certain type of disease or certain group of researchers, but rather aim to reach out to whoever needs support. Hence, when analyzing the effects of the CTSA award, it would be reasonable to consider all the projects conducted within the institution or all researchers affiliated with the institution.

2.3.3. *Support for translational research in other countries*

Similarly to the U.S., other countries around the world are also placing an increasing emphasis on translational research (Butler, 2008; Blümel, 2017). European countries created funding programs dedicated to translational research to improve the capabilities of the research system (Blümel, 2017). For instance, in the UK there has been an increase in the budget of the Medical Research Council (MRC) in 2010 with the aims to increase support for translational research (Medical Research Councils UK, 2017). This was a part of the MRC's Translational Research Strategy that was announced in the same year (Blümel, 2017; Lander & Atkinson-Grosjean, 2011). Similarly, there is a fund named Scottish Universities Life Sciences Alliance

[10] For instance, Georgia CTSA (partnership between Emory University, Georgia Institute of Technology, Morehouse School of Medicine, University of Georgia and) mentions on their website that their priorities are workforce development, collaboration/engagement, integration, methods/processes and informatics (Georgia Clinical & Translational Science Alliance, n.d.). Similarly, Northwestern University's CTSA program's three major categories of funding are pilot program, team science and career development (Northwestern University, 2019).

Assay Development Fund (SULSA) (McElroy et al., 2017). It is a £27 million fund established by the Scottish Funding Council, which has an aim to unlock the translational potential of new biology that meets industry's needs (McElroy et al., 2017).

One thing to note regarding support for translational research in Europe is that European countries conceptualize translational research differently compared to the case of U.S. (Blümel, 2017). In his study that tried to find out the differences in framing of translational research between the U.S. and European countries, Blümel (2017) maintained that the U.S. frame existing problems related to translational research as professional issues that are closely related to individual researchers whereas European countries frame the problem as an organizational issue. Blümel (2017) states that the reason for the disparity is due to the difference in their conception of science. While the U.S. emphasizes the scientific values in accordance with the scientific norms and standards, European countries focus on the problems of organization in the research sector when seeking for solutions (Blümel, 2017). In this vein, European projects with the aim to address translational research problems focus on changing the structural aspect of the research system, such as making new organizations or consortium. A representative example of this movement would the creation of European infrastructure for translational medicine (EATRIS). EATRIS, created in 2013, is a network of European biomedical translation hubs based on more than eighty academic research centers (Blümel, 2017; EATRIS ERIC, n.d.). It is the first European Research Infrastructure Consortium in the field of biomedicine (European Commission, 2017), and it provides high-end facilities and cutting-edge technologies to derisk and add value to studies related to translational research (EATRIS ERIC, n.d.).

There are similar movements in Asian countries as well. China's National Center for Translational Medicine and several other centers provide facilities and funding opportunities

specific to translational research (Williams, 2016). Also, Japan established the Translational Research Informatics Center, now the Translational Research Center for Medical Innovation, back in 2002 to facilitated the transfer of findings from basic medical research to clinical practice (Translational Research Informatics Center, 2003). It promotes academic oriented medical innovations by supporting medical researchers for all phases of biomedical study (Fukushima & Kimura, n.d.). In Korea, the concept of translational research was introduced in the 2000s and the Ministry of Health and Welfare started supporting translational research in 2005 (Kim, 2013). As various stakeholders like research universities, government subsidized research institutes, private companies and hospitals got aware of the concept, there is a growing interest in this new concept of research in Korea (Kim, 2013). However, as the definition of translational research is not consistent, similar to other countries, the Korean research community is experiencing a hard time supporting and conducting the research (Kim, 2013). However, it is apparent that the research community agrees on the importance of such type of research and believes that it will increase the chance of successful commercialization of findings from the laboratories (Kim, 2013).

CHAPTER 3. CHANGE OF COLLABORATION AMONG RESEARCH UNIVERSITIES INFLUENCED BY SUPPORT FOR TRANSLATIONAL RESEARCH

3.1. Introduction

Chapter 3 of this dissertation focuses on the change in the collaborative landscape of biomedical research among the research-intensive universities in the U.S. Promoting collaboration between institutions is one of the most important goals of CTSA award (Luke et al., 2015). NIH aims to accomplish this goal by introducing various tools that could facilitate interaction between organizations (Obeid et al., 2014).

By using co-authorship of publications as the evidence of collaboration, my analysis aims to find out if this effort leads to significant change in the network structure after the initiation of the CTSA program. Descriptive analyses of key social network analysis variables (e.g., density, centrality) were used to characterize the overall structure of the network. Then, to analyze the relationship between CTSA and frequency of interactions between academic institutions more systematically, advanced social network analysis regression model, such as Exponential Random Graph Model (ERGM) and Temporal Exponential Random Graph Model (TERGM), were used. The findings regarding the entire network showed that the biomedical research collaboration network is becoming denser and less centralized in general. For the individual institutions, the status of taking advantage of the CTSA award led to more collaborations and collaborators.

This chapter expects to contribute to the literature on the impact of the CTSA award as well as research collaboration in the field of biomedicine in general. The findings could provide insight into how the CTSA award should be operated to promote more collaboration between

institutions. Also, empirical evidence showing how and to what extent government funding changes inter-organizational collaboration patterns could help policymakers set relevant science policies more broadly.

3.2. Literature review

3.2.1. *Factors affecting research collaboration and their relevance with the CTSA award*

Collaboration in scientific research is widespread and team-based research dominates solo research in science (Walsh & Maloney, 2007; Price 1986). Scholars assert that science is a social institution where interaction between scientists is crucial to its advancement (Kats & Martin, 1997) and will contribute to technological and economic well-being (Mowery, 1998). Numerous factors exist that encourage collaboration in science, which can lead to the formation of research teams. In this section, I aim to identify the key factors that shape the collaborative landscape of biomedical research and that are related to the CTSA award. I will first point out the factors that apply to scientific research in general and try to relate them with collaborations formed due to the CTSA award.

Access to resources. One benefit of scientific collaboration is the sharing of resources (Bozeman & Corley, 2004; Katz & Martin, 1997; Gaughan, Melkers, & Welch, 2018). When scientists interact with others, they get access to the resources that they would not have access to otherwise. These resources are in various forms; tangible resources like equipment, intangible resources like professional advice (Gaughan et al., 2018). The CTSA award provides facilities with multiple use and administrative support to researchers (CTSA Principal Investigators, 2012).

As the researchers affiliated with the CTSA receiving institutions would have access to these resources, these researchers would be attractive partners for those who do not have these kinds of resource.

Financial incentive. Financial factors also affect the formation of research collaboration and this feature applies to biomedical field as well. Financial incentive is what government programs use to promote collaboration between institutions, sectors and disciplines (Defazio, Lockett, & Wright, 2009; Katz & Martin, 1997; Walsh & Maloney, 2007). In some cases, government programs provide grants only if the applicants show enough evidence of future collaboration plan in their application (Katz & Martin, 1997). This is also true in the case of the CTSA award. NIH requires CTSA applicant to provide their plan to collaborate with various stakeholders in the health community (National Institutes of Health, 2017b). Therefore, CTSA applicants are expected to seek external partners before the application and actually collaborate with other organizations during the period they receive the award.

Mentorship. Another major factor that impacts research collaboration is mentorship (Katz & Martin, 1997). With collaboration with others, scientists can gain advice and get access to the expertise of their collaborators when they get mentored (Bozeman & Corley 2004; Gaughan et al., 2018). Also, they can form an intellectual companionship with their collaborators that can be used in their future research (Katz & Martin, 1997). This is important because past collaborators may be more open to sharing their resources than new collaborators (Gaughan et al., 2018).

There is a rising demand of researcher with expertise in translational research and the CTSA program has made good progress in educating and training people on translational research (Working Group of the NCATS Advisory Council to the Director, 2014). Researchers

affiliated with CTSA receiving institutions would have a higher chance of having received training on translational research than the researchers affiliated with institutions not receiving the CTSA award. Hence, research teams involved in translational research would be seeking research collaborators who are members of CTSA receiving institutions to take advantage of their learning experience and insight in the translation process (He, Geng, & Campbell-Hunt, 2009). In this respect, we could expect that scientists belonging to the CTSA receiving institutions will be popular as research collaborators.

Source of creativity. The fact that scientific collaboration could be a source of creativity also affects scientific network formation (Katz & Martin, 1997). As scientific questions are becoming increasingly more complex (Bozeman & Corley, 2004; Gaughan et al., 2018; Katz & Martin, 1997), many problems cannot be solved with conventional intradisciplinary approach and there is more need for combination of various views (He et al., 2009; Katz & Martin, 1997). This leads to the need for collaborations that span across various disciplines (Walsh & Maloney, 2007; Lee & Bozeman, 2005; Luke et al., 2015). As collaboration with other disciplines may act as a source of stimulation and creativity, it may help the scientists come up with new techniques and subsequently lead them to save time in solving problems (Katz & Martin, 1997). This attempt to break down disciplinary silos may lead to strong scientific evidence and provide the society a great return (Luke et al., 2015).

The CTSA award promotes multidisciplinary research as mentioned in its Funding Opportunity Announcement (FOA) (National Institutes of Health, 2017b). The FOA states that the CTSA receiving institutions should provide incentives to teams with multidisciplinary members. In this regard, research teams that get support from the CTSA award are likely to try to bring researchers from various disciplines together.

However, the tools provided from the CTSA program that promote inter-organizational collaboration may have unintended consequence that further strengthens intradisciplinary collaboration. The CTSA program provides various online tools (e.g., Profiles Research Network Software) to support researchers in finding collaboration partners (CTSA Principal Investigators, 2012). These tools have the aim to help researchers locate experts in various disciplines (CTSA Principal Investigators, 2012). However, based on the claim by Van Alstyne and Brynjolfsson (1996), the information technology could lead to narrower scientific interactions. They assert that using these tools could help the researchers find people and works that really interest them, which is likely to be in their own discipline (Van Alstyne & Brynjolfsson, 1996). As a result, the phenomenon of balkanization of science by the "topic space" can be strengthened. This feature can also be applied to the web-based search tools provided by the CTSA program and more monodisciplinary research may take place rather than more multidisciplinary research.

3.2.2. *The impact of the CTSA award on research collaboration network*

Few studies have explored how NIH's support for translational research has changed the biomedical research network landscape. Luke et al. (2015) studied the case of the Institute of Clinical and Translational Sciences (ICTS) at Washington University, which received CTSA from 2007. They studied how the scientific collaborations among the ICTS members change by year. By looking at the grant application documents submitted to ICTS, they found out that researchers increase their collaboration with others in both project planning and dissemination of their research finding. This result holds for two study samples, which are the group with only the members who were ICTS members from the first place and the other group including all

additional members who joined ICTS during the period of analysis (2007 -2010). Hughes et al. (2010) looked at the case of University of Pennsylvania, another CTSA recipients. They investigated the Institute for Translational Medicine and Therapeutics (ITMAT) at University of Pennsylvania and saw how the collaborative landscape changed using social network analysis methods. By looking at co-authorship patterns, they showed that not only the network size among ITMAT's member researchers grew substantially but also the frequency of collaboration has increased dramatically. The network structure within University of Arkansas, which started receiving the CTSA award in 2009, was analyzed by Bian et al. (2014). By looking at the change of collaboration network in research grants of University of Arkansas for Medical Sciences before and after receiving the CTSA award, they found out that research teams are becoming more multidisciplinary.

These studies looked at the impact of CTSA mainly using descriptive analysis on a single or small group of organizations. They show that the CTSA award clearly had an impact on the intra-organizational research collaboration. However, they did not investigate inter-organizational collaboration change, which is one of the core goals of the CTSA award. Therefore, it would be meaningful to look at the impact of the CTSA award on the collaboration network among institutions. By doing so, we could be able to figure out whether the tools[11] that were introduced by CTSA institutions with the purpose to promote inter-organizational collaboration were effective or not. We could also see if various factors affecting research collaboration, which are mentioned in the previous section, come into play in shaping the network.

[11] One example is the electronic research network system that provides information on researcher's status and ongoing projects (Obeid et al., 2014). From the survey of CTSA member institutions, Obeid et al. (2014) found that this system was effective in fostering cross-institutional collaboration.

3.3. Theoretical background and hypotheses

Scientists choose to collaborate depending on a variety of strategies and motives (Lee & Bozeman, 2005; Bozeman & Corley, 2004) and various theories can be applied in explaining the factors related to biomedical research collaboration. The first group of theories I want to mention is related to resource dependence approach. These resource-based theories can be used to explain the phenomena that the ones with more resources tend to have more links than the ones with fewer resources (Katz, Lazer, Arrow, & Contractor, 2004). This theory explains the general phenomena that research groups with large grants tend to have many collaborators (Bozeman & Corley, 2004). One of the theories adopting this approach is the theory of social exchange and dependency. This theory claims that people tend to form ties with others that they can exchange valuable resources with or get valuable resources from (Homans, 1950). As scientific projects are becoming more difficult and the structural complexity of collaboration increase (Walsh & Maloney, 2007), scientists need more resources for the successful execution of projects. Therefore, the ones that lack resources (e.g., supporting facilities, knowledge on the field) would try to form a relationship with others for the successful execution of a project (Walsh & Maloney, 2007).

As noted, the CTSA program provides resources for instrumentation, collaboration, training, pilot research and administrative support for translational research (Clinical and Translational Science Awards Consortium, n.d.-a; Leshner et al., 2013). Hence, based on the social exchange theory, the researchers affiliated with institutions having scarce resources, in this case those not taking advantage of the CTSA award, would be willing to make ties with CTSA recipients as they can benefit from the resources that they do not possess. This would lead to the situation where the researchers experience an increase of inter-organizational collaborations with other

researchers after their affiliated institutions start taking advantage of the CTSA award. The CTSA receiving institutions would also want to form collaborations with institutions not receiving the CTSA award based on the self-interest paradigm, which will be discussed later. To summarize, we can expect that the CTSA recipients will have more collaborators than the institutions not receiving CTSA based on the resource-based theory and I set the following hypotheses based on this expectation.

***Hypothesis 1*: The institutions receiving the CTSA award would have more collaborating institutions than the institutions not receiving the CTSA award when the impact of the CTSA award gets realized.**

***Hypothesis 2:* The institutions receiving CTSA award would experience higher growth in the frequency of collaborations with their partners than the institutions not receiving the CTSA awards when the impact of the CTSA award gets realized.**

Hypothesis 1 is related to the increase of collaborating partners following the receipt of the CTSA award. A binary measure on whether you have a connection with others will be used to test this hypothesis. On the other hand, hypothesis 2 is related to the frequency of interaction between institutions. Therefore, a weighted measure incorporating the frequency of collaborations will be used for testing the hypothesis. When testing either of the hypotheses, we need to take into account the time lag between funding and the point when the impact of the award gets realized.

The other theory related to the resource-based approach in collaboration is the theory of self-interest paradigm. This theory explains that a tie is formed between two parties if it increases

each individual's interest (Katz et al., 2004). The common interest of those who apply to a grant pursue is increasing their chance of being selected as the recipient of the grant. In that regard, in case the funder of a grant states that collaborating with certain types of partners would give them a higher chance of acceptance, potential applicants would seek for that kind of partners for the application. In the case of CTSA program, with the aim to create regional hubs for translational research, its FOA (National Institutes of Health, 2017b) requires institutions to provide plans to collaborate with a diverse range of partners including local communities, government, non-profit organizations and other universities. Therefore, to raise their chance of receiving the grant, potential applicants of CTSA would try to make partnership with these parties before they submit their applications to NIH. In this respect, self-interest approach can help us understand why CTSA institutions reach out to other institutions, which in turn will bring CTSA institutions more collaborators in the periphery. Based on this expectation, I developed the following hypothesis.

Hypothesis 3: **The biomedical research network will become less centralized as the institutions taking advantage of the CTSA award interact with institutions not taking advantage of the CTSA award more than before.**

3.4. Data and methodology

3.4.1. *Empirical setting*

Co-authorship as the evidence of collaboration

Selecting the best indicator to measure research collaboration is a challenge (Katz & Martin, 1997; Lee & Bozeman, 2005). As Katz and Martin (1997) noted, research collaboration

is an ill-defined field as there are a variety of views on the concept of research collaboration. Researchers have been trying to find out the most appropriate measure for research collaboration and co-authorship is one of the most widely used measures in measuring scientific collaboration (Balconi et al., 2004, Katz & Martin, 1997; Lander 2013; Newman 2004). Co-authorship is an undirected measure that takes individual authors or author's affiliations as nodes and existence of coauthored paper as links (Ding, 2011).

When doing research on scientific research network, using co-authorship has many advantages. Balconi et al. (2004) claimed that it is an ideal quantitative indicator that examines the relationship between academic scientists if knowledge exchange occurs in team-working experience, which is the case in most collaborative scientific research. Co-authorship is generally regarded as an unobtrusive measure of research collaboration (e.g., Qin, Lancaster, & Allen, 1997) with stability over time and good verifiability (Katz & Martin, 1997). It can also be transformed into other measures like the number of authors, the number of institutional affiliations and the number of disciplines (Qin et al., 1997) and other social network analysis measures.

It is true that using co-authorship as the measure of collaboration has some limitations. First, co-authorship may not always mean that actual collaboration took place. For instance, people who just provided materials, the "guest authors," or those who just did routine tasks may also be listed as one of the coauthors (Stokes & Hartley, 1989; Jabbehdari & Walsh, 2017). However, there are clear criteria for authorship for journals in life science, which is focus of my study. For instance, International Committee of Medical Journal Editors (ICMJE) has its own guideline on authorship and it is adopted by 600 biomedical journals (Jabbehdari & Walsh, 2017). Their guideline states that to be listed in the coauthor list a person should have substantial

contribution to the work, should have taken critical role in drafting or revising the work and should have participated in the final approval of the work (International Committee of Medical Journal Editors, 2019). This kind of guidelines could reduce the concern about the existence of the "guest authors" and increase the reliability of the co-authorship data. Secondly, co-authorship can only capture official collaborations. Research collaboration can take various forms that don't leave any records such as discussions in meetings. By nature, these kinds of informal collaborations cannot be captured by co-authorship. However, we should note that all other indicators, not only the indicators based on co-authorship, cannot capture unofficial collaboration easily. Very time-consuming job, like interviewing all researchers, would be needed to capture undisclosed relationship regarding research collaboration. Given the scope of my research, using this kind of approach in measuring collaboration is almost impossible. Hence, even using co-authorship has some limitations I will be using this "codified markers of collaboration" (Park, Yoon, & Leydesdorff, 2016, p. 1017) as the evidence of collaboration in my study.

Study sample

In this study, the population of interest is 115 Carnegie Highest Research Universities (Carnegie R1 universities) in The Carnegie Classification of Institutions (Indiana University Center for Postsecondary Research, n.d.). Two factors led me to choose R1 universities as the study sample. The first factor is the size of the treated group and the controlled group. As seen in Table 1, there are 47 universities in the R1 university group that have experience of receiving the CTSA award. In addition, there are eight other R1- universities that have official partnership with the CTSA receiving institutions. This leaves us with 60 institutions that are not affected by the CTSA award, which can be used as the comparison group. This means there are almost the

same numbers of institutions in both the treatment group and the control group if we take R1 universities as the study sample.

Table 1. Institutions classified by leadership in the CTSA program (as of 2015)

Group	Lead CTSA institution	Non-lead CTSA institution	Non-CTSA institution	Total
R1 Universities	47 (40.9%)	11 (7.0%)	60 (52.2%)	115 (100%)
R2 Universities	5 (4.7%)	5 (4.7%)	97 (90.7%)	107 (100%)

Note: The full list of Carnegie R1 universities in each leadership type can be found in Appendix A

The second and more important factor that led me to take R1-university group for the analysis is related to the comparability of the treatment and the control group. To have high level of internal validity of research, we need to have appropriate comparison group that has similar features with the treatment group (Campbell & Stanley, 1963). The most preferred case would be introducing a control group that has same attributes with the treatment group before the treatment takes place. However, it is impossible to have that kind of pair in a quasi-experimental setting like my study. Hence, in this situation, I should try to get the control group that is as similar as possible to the treatment group.

As expected, there is a clear difference in the values regarding R&D attributes for R1 universities and High Research Activity University (R2 universities) (Table 2). For instance, average R&D expenditures in Science and Engineering (S&E) for R1 universities are 6.6 times larger than that of R2 universities. Also, per-capita S&E research staff is four times larger for R1 universities than R2 universities. This suggests that R1 universities and R2 universities are fundamentally different in terms of the research capacity. If I included R2-universities in the study sample, I would be comparing a large number of R1 universities receiving the CTSA

award to a large number of R2 universities not receiving the CTSA award. This will lead to a large bias of the result as these two groups might have differed even without the impact of the CTSA award. This would lower the internal validity of the research and limit the strength of the claims I could make on the causal relationship between the CTSA award and the dependent variables of my analyses.

Table 2. Mean values of R&D related attributes for R1 and R2 universities in 2014

Group	S&E R&D Expenditures*	S&E Research Staff	STEM Doctorates	Per-capita S&E R&D Expenditures*	Per-capita S&E Research Staff
R1 Universities	411,742	604	202	277.9	0.4
R2 Universities	62,821	51	37	144.7	0.1

Note: * indicates units in $1,000

Source: Indiana University Center for Postsecondary Research (n.d.)

Including only R1 universities in my sample means that I aim to look at the collaboration between universities that have high research ability, which was the approach Owen-Smith and Powell (2003) chose in their study on patenting activities of universities. Regarding the field of science, I restricted it to life science and related fields. To do so, I only looked at the grants from NIH and their outcomes. Using outcomes from grants has the advantage of capturing the dissemination of knowledge and findings from grant-supported projects (Ihli, 2016). For the period of analysis, I chose the years between 2001 and 2015. This means that I have enough years before the initiation of CTSA, which is five years, and after the initiation of CTSA, which is nine years. The organizations that received CTSA after the first year of the start of CTSA would have fewer years after the event but more years before the event. Therefore, for some analyses, I normalize the period taking the year of designation as the base year. I didn't use the

most recent years (e.g., 2016 and 2017), in my analysis because there should be some time lag for the outcome of projects to be realized. Also, it takes some time to capture all outcomes (e.g., publications from NIH projects) in the NIH reporting system.

The status of being impacted by the CTSA award will be the criteria for dividing the treatment group and the control group. The institutions that received CTSA award will be in the treatment group whereas the institutions that didn't will be members of the control group. One issue to address in identifying the institutions in the treatment group is the appropriate classification of partnering institutions of the CTSA award receiving institutions. NIH promotes CTSA receiving institutions to partner with other institutions and encourages them to share their resources with their partners (Llewellyn et al., 2019). In this regard, some CTSA receiving institutions make official partnership with other institutions with the purpose to make the best use of the award, which make them multi-hub institutions (Llewellyn et al., 2019). As the partnering institutions also have access to the resources, methods, and other tools that could accelerate the research process (Llewellyn et al., 2019), I also include the partnering institutions in the treated group. The basic rules that were applied in identifying the partnering institutions are as follows.

- Classify an institution as CTSA-partner if that institution never received the CTSA award but made official CTSA partnership with any of the CTSA receiving institution (CTSA-leader) that would not have been formed without the award.
- The CTSA partnership is considered official if 1) CTSA-leader institutions self-identified partnership with CTSA-partner on their hub website or their grant abstracts, 2) there is

leadership or principal investigator from both CTSA-leader and CTSA-partner listed on their hub website or their grant abstracts

- The universities that would have been connected even without the effect of the CTSA award (e.g., universities in same state university system) were not considered to have an official CTSA partnership

Figure 8 shows the change in the size of the treated group and the controlled group by year[12]. We can note that the size of both the treatment group and the control group will change by year. Before the CTSA program started (e.g., year 2003), all Carnegie R1 universities are classified as a member of the control group. The number of institutions in the treatment group increases as the CTSA program progresses (e.g., 35 in 2008 and 55 in 2015).

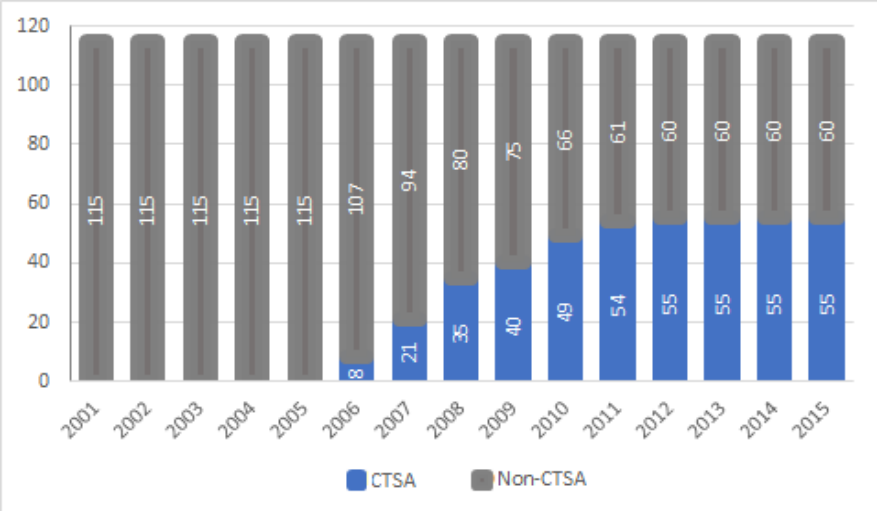


Figure 8. Change of the sizes of the treatment group and the control group by year

[12] Full list of institutions by CTSA-leader type can be found in the Appendix A.

Data construction

The primary data source for my research is from NIH Research Portfolio Online Reporting Tools[13] (RePORTER). As part of NIH's "open government" initiative and Public Access Policy, NIH provides information on all projects it funded through RePORTER (National Institutes of Health, 2018; Surkis et al., 2016). CTSA receiving institutions also have obligations to report publications that received support from the CTSA program's services, resources, training programs and pilot projects (Surkis et al., 2016). I used ExPORTER (National Institutes of Health, 2017a)[14], which is part of the RePORTER system that contains downloadable version of the data, to download the dataset I needed. I first downloaded them in Comma Separated Value (CSV) format and transformed the data into the format that can be imported into the statistical programs I used for my analysis. The dataset from ExPORTER has a number of variables related to the projects including core project number, the fiscal year of the project, total cost, support year, administering institution, affiliation of principal investigator and so forth. I used core project number to link the information of the projects with the publication outcomes.

To get the collaboration network between organizations multiple steps had to be taken. The first step is getting the list of projects that are conducted by Carnegie R1 universities between 2001 and 2015 (Carnegie R1 Projects). To do this, I download CSV files that contain information on projects supported by NIH from NIH ExPORTER (National Institutes of Health, 2017a). Then, I search for projects where the principal investigator (PI)'s affiliation was one of the 115 Carnegie R1 universities. I found out there were 139,066 projects, which is the count of unique core project number throughout the period.

[13] Source: https://projectreporter.nih.gov/reporter.cfm
[14] Source: https://exporter.nih.gov/ExPORTER_Catalog.aspx

The second step is extracting journal articles that are generated from Carnegie R1 Projects (Carnegie R1 Publications). For this task, I downloaded another set of CSV files from NIH ExPORTER (National Institutes of Health, 2017a) that contain PubMed IDs of articles in one column and acknowledged NIH projects in another column. Using this file, I identified and extracted a group of matches between PubMed ID and NIH core project number. I found out that there were 840,464 publications with Digital Object Identifier (DOI), which means 6.04 publications with DOI were produced per project on average. Here, we need to keep in mind that only the articles that cite NIH projects properly can be captured in this dataset (Han et al., 2018). Articles that do not acknowledge any grant number or acknowledge a different grant number, whether it was intentional (e.g., for the better assessment of specific project) or unintentional (e.g., by mistake), could not be properly included in the dataset.

As the third step, bibliometric information of Carnegie R1 Publications is downloaded from Web of Science (WOS) via access through Georgia Tech Library. WOS is chosen as it contains comprehensive data of nearly 12,000 high impact journals (Lander, 2013; Clarivate Analytics, 2018) and it covers journals in all fields of interest (e.g., science, social science and arts and humanities) for over 100 years (Clarivate Analytics, 2018). In addition, it offers citation indexes that can be used in various applications, which will also be used in the later section of this dissertation.

As the fourth step, I identified organizations that participated in the publication of the articles (Carnegie R1 co-authoring organizations). I assumed that if the researchers appeared in a coauthor list of a same publication, the organizations that these researchers are affiliated with collaborated[15], which is the approach that studies looking at inter-organizational collaboration

[15] One issue of this approach is that if an author lists two or more organizations as his or her affiliations, those organizations would be considered as collaborating institutions. In this dissertation, I do count them as collaboration

use frequently (e.g., Lander, 2013). In this step, different ways of referring to the same organization need to be merged. For example, Columbia University and Columbia University in New York City must be merged into Columbia University. This job of merging institution names was conducted using self-made thesaurus for use in VantagePoint.

For the last step, I made the collaboration network composed of only 115 Carnegie R1 universities. I created this network by dropping non-Carnegie R1 universities from the network matrix that is composed of all Carnegie R1 universities and Carnegie R1 co-authoring organizations. The final outcome is the collaboration network matrix of 115 Carnegie R1 universities, which will be used as the dependent variable in my analysis. Figure 9 illustrates the flow of the five-step data making process.

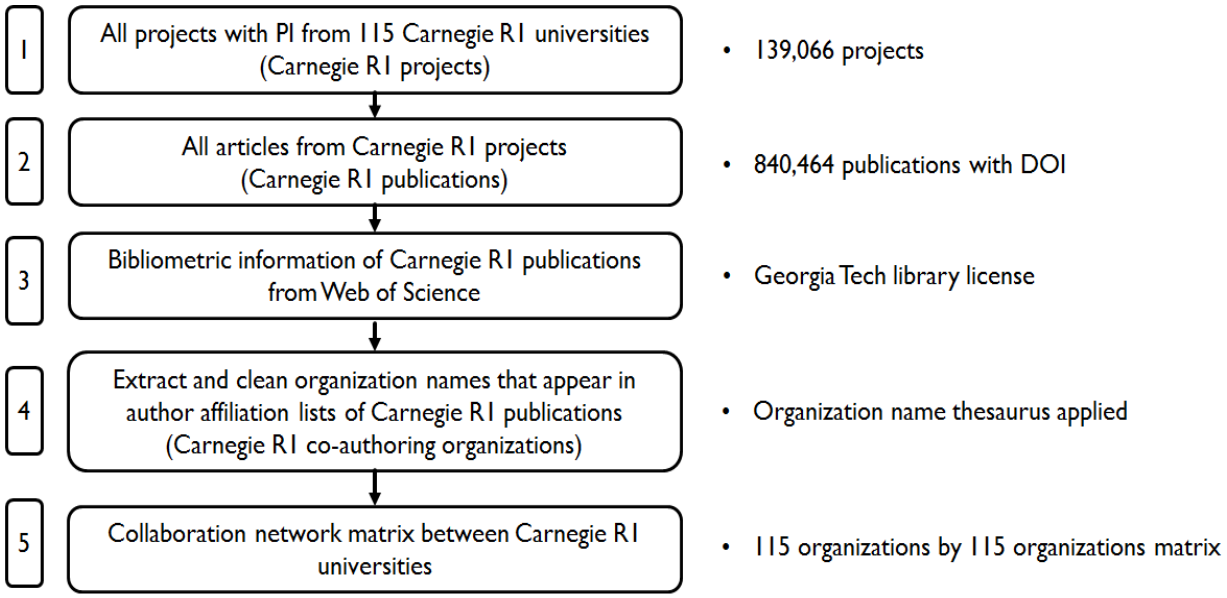


Figure 9. Flow chart of making collaboration network between Carnegie R1 universities

institutions as I agreed with the claim by Lander and Atkinson-Grosjean (2011) and Swan et al. (2007). They asserted that this kind of scientists are 'boundary spanners' who take the role of sharing knowledge between their affiliated organizations.

The programs I used for my analyses are Vantage Point, R, STATA and Microsoft Excel. With Vantage Point, I extracted bibliometric information that is needed for the analyses. The factors that were extracted are DOI, author affiliation, publication year and Web of Science Category. Furthermore, Vantage Point's co-occurrence matrix function was used to create network matrices. In some cases, I created thesauruses that can be used in jobs related to classifications and merging observations. R was used to construct social network data, calculate social network measures (e.g., density, centralization, centrality) and run network regression analyses. For some descriptive analyses and regressions, STATA was used. Microsoft Excel was used to get basic statistics and for some data visualization.

3.4.2. *Measures and variables*

Social network analysis (SNA) provides us with insights on the relationship between multiple individuals or entities (Gaughan et al., 2018). The scientific relationship is one of the topics that the SNA approach is well-suited to analyze (Katz et al., 2004; Luke et al., 2015)

Macro-level network measures

Network density. Network density is calculated by the number of actual ties divided by the number of possible ties (Hanneman & Riddle, 2005). The value is same as the ratio of existing adjacencies divided by the number of possible pairs in the network (Hanneman & Riddle, 2005). The density gives us insight on how close the entities in the network are to each other and the value is associated with the speed of information diffusion in the network (Hanneman & Riddle, 2005).

Network centralization. Many variables provide a general picture of a network and centralization is of one of them. Centralization measures how focused a network is on few core nodes (Haythornthwaite, 1996). The network with high value of centralization represents highly centralized network, which means that there is small number of core nodes that are connected to a large portion of other nodes (Haythornthwaite, 1996). In this case, these core nodes act as the hubs of the whole network and they have more power compared to non-hub nodes (Hanneman & Riddle, 2005).

Micro-level network measures

Centrality measures. Four centrality measures, which are degree centrality, closeness centrality, betweenness centrality and eigenvector centrality, are the variables that I will use in my descriptive analysis. These measures provide us information on key features of the collaborative landscape (Hanneman & Riddle, 2005, Katz et al., 2004). For instance, with these measures, I would be able to figure out how central an institution is in the collaborative network and get a sense of their position in the network (Hanneman & Riddle, 2005). Table 3 provides the concept and features of four centrality measures I used.

Table 3. Concepts of centrality measures

Variables	Concept and features
Degree Centrality	· How many institutions that an institution has direct collaboration with · Large value represents high prestige and prominence
Closeness Centrality	· The extent to which an institution is directly or indirectly connected to the rest of the nodes in the network · Measured using the sum of distance to others from an ego and it indicates how fast can an institution reach others in the network
Betweenness Centrality	· How much an institution has links with others who are not directly connected and become the direct route between two alters in the network · High value means less dependency on others and having more power
Eigenvector Centrality	· How well an institution is connected to other well-connected institutions · Focusses on the global structure of the network rather than the local structure of the network

Note: Descriptions mainly from Hanneman & Riddle (2005)

Variables for regression models

Independent variable. As I am interested in the impact of receiving CTSA on collaboration formation and change, I used the dummy variable related to receiving CTSA as my key independent variable and called it **CTSA**. The number of CTSA receiving Carnegie R1 universities rose from eight in 2006 to fifty five in 2016. Therefore, the binary independent variable $\boldsymbol{x}_{i|j(t)}$ will have the value of 0 for all organizations till 2005 and eight organizations will start to have the value of 1 from 2006. The number of organizations having the value of 1 for this variable will increase to 21 in 2007 and reach its highest value of 55 in year 2012. The full list of the CTSA receiving institutions can be found in the Appendix A.

Control variables. Various attributes of the organizations were used as the control variables, which is denoted as $\boldsymbol{S}_{ij(t)}$. First, from the Carnegie Classification of Institutions of Higher Education website (Indiana University Center for Postsecondary Research, n.d.) whether an organization has a medical school (**Medical**), whether an orgnization is controlled by a private party (**Private**) and the state of the organization (**State**) is obtained. Secondly, R&D expenses in the life sciences (**R&D**) were obtaned from National Science Foundation (NSF) Higher Education Research and Development Survey (National Science Foundation, n.d.-b). The field selected to get this value was "Life sciences, all." and I took the sum values of all affiliated institutions of the corresponding universities. Thirdly, NSF Survey of Graduate Students and Post-doctorates in Science and Engineering (GSS) (National Science Foundation, n.d.-c) was used to calculate the number of researchers in life science for each organization (**Researcher**). With the raw data from NSF GSS, I added the number of graduate students, post doctoral

students and non-faculty researchers in life sciences[16] to get the value of a total number of researchers in life science. Table 4 shows the list of control variables and how they were coded.

Table 4. List of control variables used in the analytic model

Name	Code	Source
Medical	0: Does not grant medical degrees 1: Grants medical degrees	Carnegie Classification of Institutions of Higher Education website[17]
Private	0: Public university 1: Private university	
State	1: AL, 2: AK, 3: AZ, 4: AR, 5: CA, 6: CO, 7: CT, 8: DE, 9: DC, 10: FL, 11: GA, 12: HI, 13: ID, 14: IL, 15: IN, 16: IA, 17: KS, 18: KY, 19: LA, 20: ME, 21: MD, 22: MA, 23: MI, 24: MN, 25: MS, 26: MO, 27: MT, 28: NE, 29: NV, 30: NH, 31: NJ, 32: NM, 33: NY, 34: NC, 35: ND, 36: OH, 37: OK, 38: OR, 39: PA, 40: RI, 41: SC, 42: SD, 43: TN, 44: TX, 45: UT, 46: VT, 47: VA, 48: WA, 49: WV, 50: WI, 51: WY	
R&D	R&D expenditures in life sciences	NSF Higher Education Research and Development Survey (HERD)[18]
Researcher	Number of graduate students, post-docs and non-faculty researchers in life science	NSF Survey of Graduate Students and Post-doctorates in S&E (GSS)[19]

[16] Fields included are biomedical engineering, agricultural sciences, biological sciences, health science and neuroscience.
[17] Source: http://carnegieclassifications.iu.edu/downloads.php
[18] Source: https://www.nsf.gov/statistics/herd/pub_data.cfm
[19] Source: https://www.nsf.gov/statistics/srvygradpostdoc/pub_data.cfm

3.4.3. *Models*

To analyze the relationship between the variables and collaboration formation between institutions, I used the Exponential Random Graph Model (ERGM), which is widely used approach on statistical modeling of networks (Robins, Pattison, Kalish, & Lusher, 2007; Schaefer & Marcum, 2017). ERGM assumes that a given network is one possible form of random network and provides the probability of a tie conditional on all other possible networks (Schaefer & Marcum, 2017). ERGM derives parameter values that could reproduce the given network assuming that the network is at equilibrium (Schaefer & Marcum, 2017). EGRM is a cross-sectional model, which mean that it can only deal with variables for a single point of time (Leifeld, Crammer, & Desmarais, 2018)

I ran a total of 15 ERGM models, one for every year between 2001 and 2015. With these 15 models, I checked the change of the coefficients over time. The equation of the ERGM model using weighted values for the collaboration is as follows.

$$w_{ij} \sim Poisson(\mu_{ij}) \quad (1)$$

$$\log\left(\mu_{ij(t+3)} \middle| G^{c}_{ij(t+3)}\right) = \beta_1 + \beta_2 \boldsymbol{x}_{i|j(t)} + \beta_3 \boldsymbol{S}_{ij(t)} \quad (2)$$

, where *i* is a receiver effect and *j* is a sender effect.

The left side of the equation measures the rate of collaboration between *i* and *j* for year t+3. Year t+3, not t, is used to account for the time between receipt of the CTSA award and when resulting changes might be observed. I assume that it takes about three years between the time an institution receives the CTSA award and the appearance of any visible changes. This assumption is based on the finding by Ihli (2016). In her study on the lag time between the award of the grant from NSF and publication of the articles acknowledging the grant, which is called

acknowledgment lag, Ihli (2016) found out that highest count of the lag-time was three years. Though there would be some differences between NSF grant and NIH grant, I assume that they would share many features and believe it is reasonable to take this approach as most characteristics of this lag would not differ much between two government funding agencies.

One advantage of the ERGM is that it can incorporate homophily effect in the model. Homophily effect explains the tendency of a one forming a tie with the ones with similar characteristics (Katz et al., 2004; McPherson, Smith-Lovin, & Cook, 2001; McFarland, Moody, Diehl, Smith, & Thomas, 2014). This results in individuals with similar attributes interacting with each other more often than with other individuals without similarity (McFarland et al., 2014). Homophily effect has been widely studied in various topics such as friendship, advice, work, support and information transfer (McPherson et al., 2001). In the context of biomedical research, we can expect that the scientists affiliated with institutions that receive grants with similar attributes will tend to collaborate with each other more so than with the ones not receiving them. In this respect, the homophily effect of the CTSA award will also be tested in the model.

We can also expect that the organizations located in similar regions, or close-by regions, will tend to collaborate frequently. This claim can also be linked to the relationship between geographical proximity and rate of collaboration. Many studies found out that people tend to work with the ones who are physically close to them (e.g., Bozeman & Corley, 2004). Bozeman and Corley (2004) found out that researchers tend to work frequently with the people in their own working group. This could be linked to the claim that the ones in the same region will collaborate more frequently than the ones not in the same region, which will also apply to

biomedical research. Hence, I also plan to test the homophily of regions, which is analyzed in state level.

Beyond the analysis with cross-sectional ERGM, I also test the relation between variables using the temporal network regression model. To analyze across-time change of the network formation, I used Temporal Exponential Random Graph Model (TERGM). TERGM is an extension of the ERGM that account for temporal characteristics of the network formation and it can take care of time dependence over a series of networks (Leifeld et al., 2018). This model divides the dependencies into within-time change and across-time change (Shaefer & Marcum, 2017). Regarding the across-time change, the model tests the effect of the past network on the current network (Shaefer and Marcum, 2017; Leifeld et al., 2018). The model can also test whether there is a dependence between ties over time, which is called "memory terms" (Shaefer and Marcum, 2017; Leifeld et al., 2018, p. 4). These terms will be described more in detail in the result section with interpretation on actual coefficients. One limitation of TERGM is that it can only account for binary feature of the links as social network packages have not gotten to the level to deal with weighted outcomes. This means that the dependent variable is based on the existence of the link between the institutions, not the frequency of the collaboration.

3.5. Results

3.5.1. *Descriptive analysis*

Macro-level network measures

To analyze change of the overall features of the whole collaboration network among all 115 Carnegie R1 universities, I checked the trend of density and centralization. As shown in

Figure 10, the density measure of binary network increased from 0.412 in 2001 to 0.820 in 2015 showing an overall increasing trend. It means that in 41.2% of possible ties between Carnegie R1 universities were present and the portion increased to 82.0% in 2015. This means that Carnegie R1 universities tend to collaborate more with each other over time.

On the other hand, the centralization measure showed a decreasing trend, and the value dropped from 0.446 in 2001 to 0.183 in 2015. As the centralization measure is in a decreasing trend, we can claim that the whole network is becoming less centralized and this confirms that my hypothesis 3 holds. This means that the number of universities acting as the hub of the network is increasing. In the early years, there may have been a small number of Carnegie R1 universities monopolizing the collaborations with other Carnegie R1 universities. However, in later year the number of institutions that act as hubs, with a lot of cooperation with other Carnegie R1 universities, has increased.

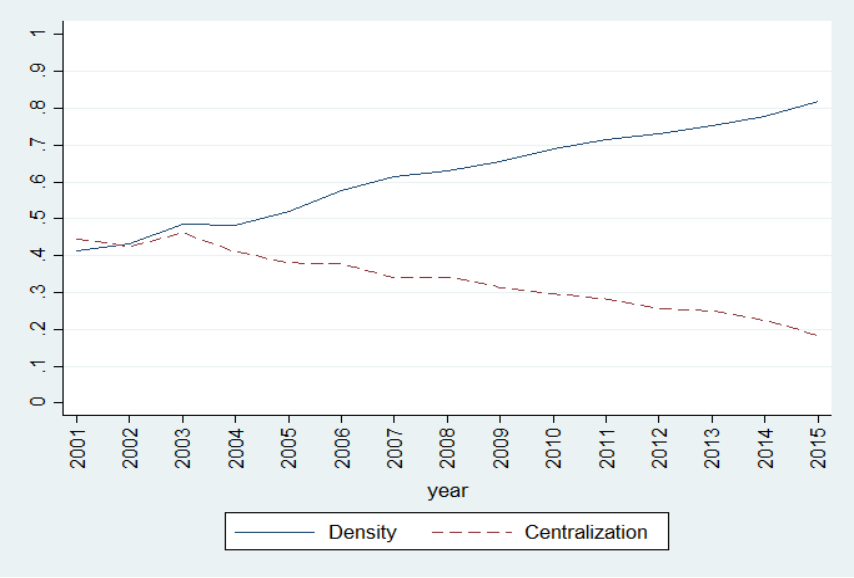


Figure 10. The trend of density and centralization of the whole network

To figure out the reason that led to this change, I divided the whole network into 1) a sub-network that is only composed of the institutions that ever took advantage of the CTSA award (CTSA-evers) and 2) another sub-network that is only composed of the institutions that never took advantage of the CTSA award (CTSA-nevers), which are the institutions that stay in the control group until 2015. Figure 11 and Figure 12 show the change of density and centralization of two sub-networks. For the first sub-network, we can see that there were slight changes in the density. The change for this variable was only around 0.10 during the whole period between 2001 and 2015. On the other hand, if we check the change of the density for the second sub-network, we can see that there was a substantial increase, from 0.15 in 2001 to 0.60 in 2015. These findings indicate that the increase of density is due to both the increase of collaboration among institutions in CTSA-ever group and collaboration among CTSA-never institutions, but the change is mainly due to the increase of collaboration among CTSA-never institutions. It can be seen that CTSA-never institutions have a larger increase in the number of collaborations than the CTSA-ever group.

If we look at the centralization measure, we can see that the measure is not decreasing in the second sub-network, but decreasing in the first sub-network. Therefore, we can claim that the decrease of centralization in the whole network is mainly due to the decrease of centralization between institutions that took advantage of the CTSA award. From the finding, we can expect that the network hubs are formed in the CTSA-ever group.

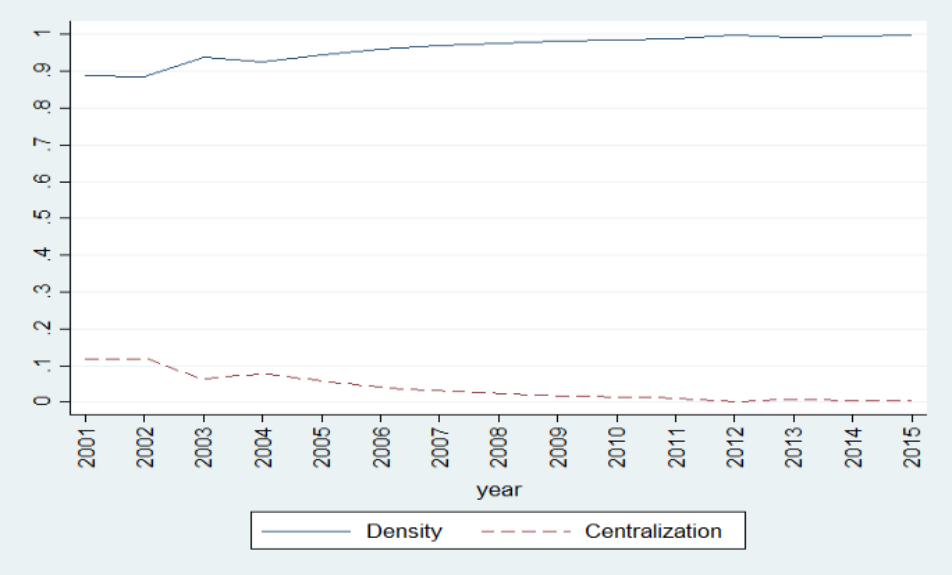


Figure 11. Centralization and density for the network between CTSA-ever institutions

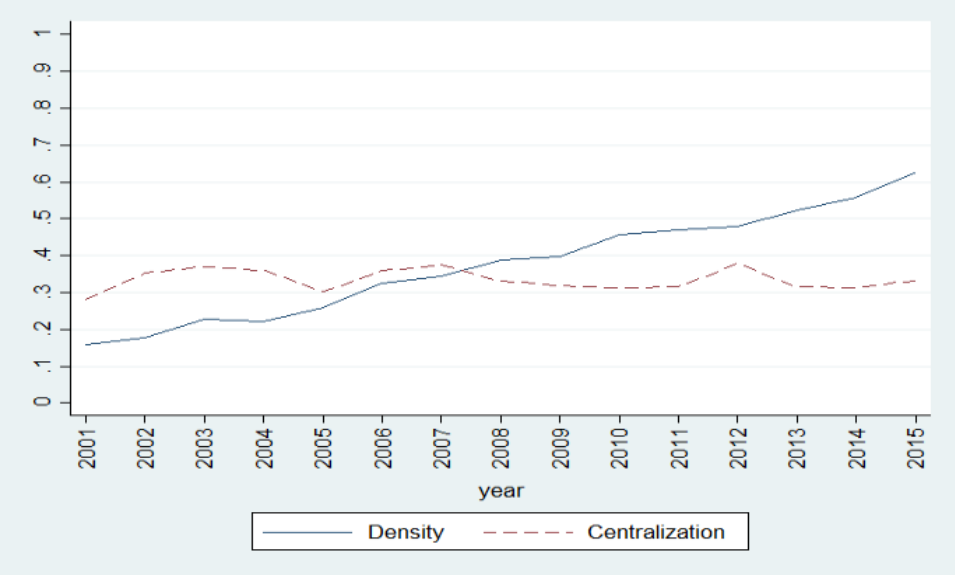


Figure 12. Centralization and density for the network between CTSA-never institutions

To investigate this issue further, I looked at the frequency and portion of collaborations having CTSA-ever group institution as partners or CTSA-never group as partners. The trend of

frequency having calendar years as the x-axis are shown in Figure 13 and Figure 14[20]. Figure 15 and Figure 16 shows the trend of the portion of collaborations. The solid line and dashed lines represent the mean values and the shaded areas show the range between 25 percentile value and 75 percentile value. The vertical line shows the year when the CTSA program started, which is 2006. From the figure, we can see that the collaboration frequency with CTSA-ever institutions is increasing at a faster rate than the collaboration frequency with CTSA-never institutions. We can note that even the CTSA-never group institution is collaborating more with CTSA-ever group institutions and CTSA-ever group institutions barely collaborative with CTSA-never group institutions. The portion of collaboration with CTSA-never group institutions is less than 20% of the total collaboration that CTSA-ever institutions have.

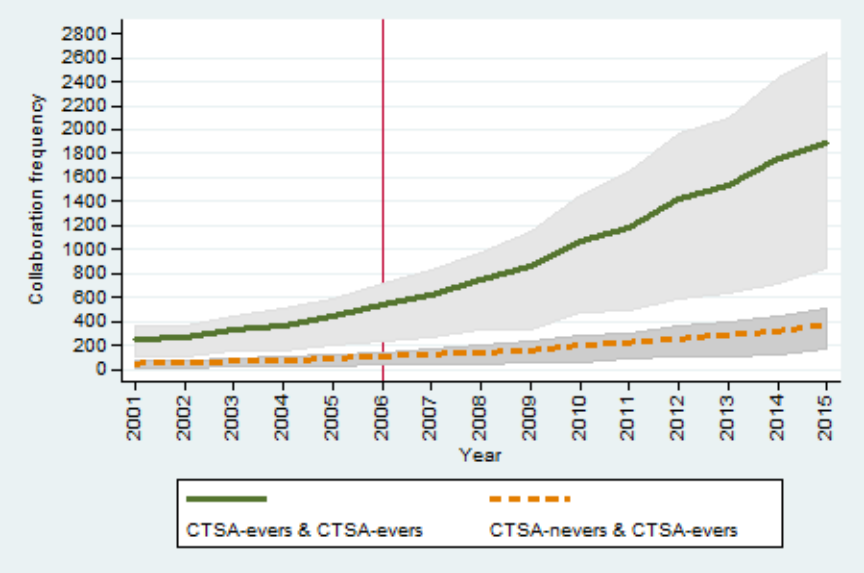


Figure 13. Collaboration frequency with CTSA-evers

[20] The graphs with normalized years as the x-axis are in the Appendix B. These graphs show similar pictures with the graph having calendar year as the x-axis.

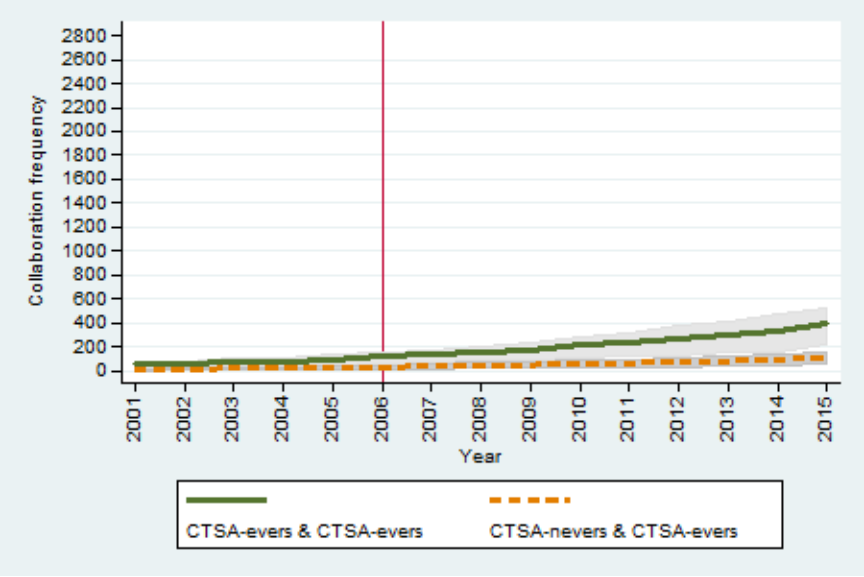


Figure 14. Collaboration frequency with CTSA-nevers

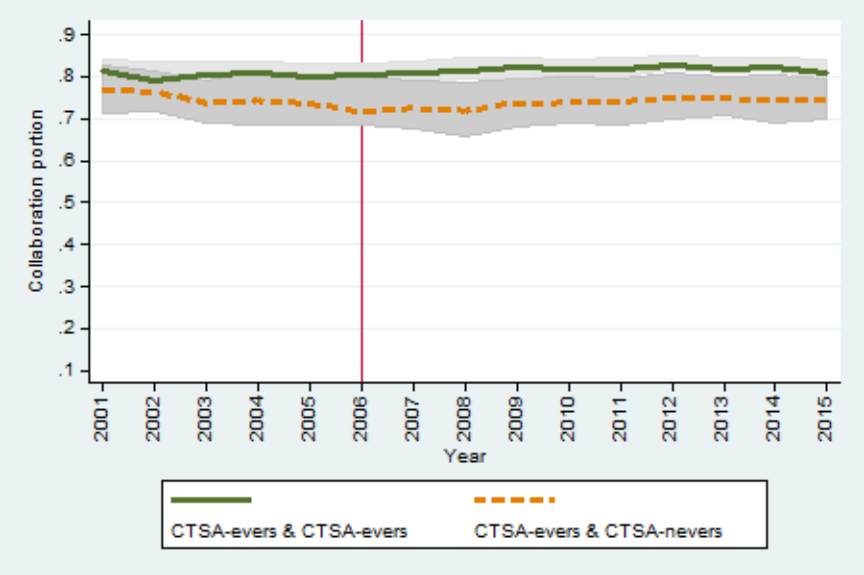


Figure 15. Collaboration portion with CTSA-evers

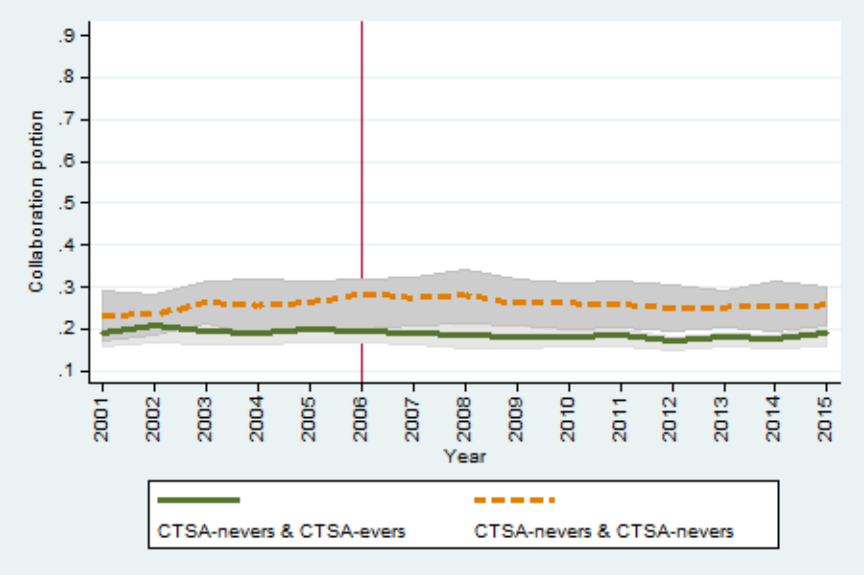


Figure 16. Collaboration portion with CTSA-nevers

Micro-level network measures

Figure 17 through Figure 20 show the trends of four kinds of centrality measures, which are degree centrality, closeness centrality, betweenness centrality and eigen-vector centrality, of institutions in CTSA-ever institutions and CTSA-never institutions[21]. Degree centrality and closeness centrality increased for both treatment group (CTSA-ever institutions) and control group (CTSA-never institutions). This result indicates that institutions in both the treatment group and the control group experienced an increase in the number of collaborations, as expected.

If we look at the change in the betweenness centrality, we can see that there is a decreasing trend for the treatment group whereas the institutions in the control group are not experiencing any big change. If we look at eigen-vector centrality value, there is a slightly decreasing trend for the CTSA-ever group and there is clear increasing trend for the CTSA-never group. This means that institutions that took advantage of the CTSA award are increasing their

[21] The graphs with normalized-years as the x-axis are in the Appendix B.

number of collaborations with the institutions in the periphery. On the other hand, the CTSA-never group institutions are getting more connected with the institutions that have many links of its own.

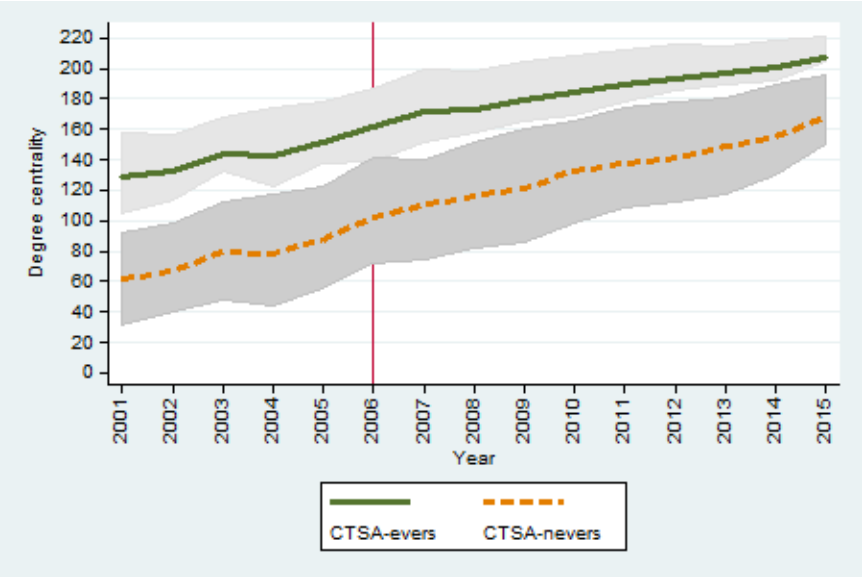


Figure 17. Degree centrality trend of institutions in CTSA-ever group and CTSA-never group

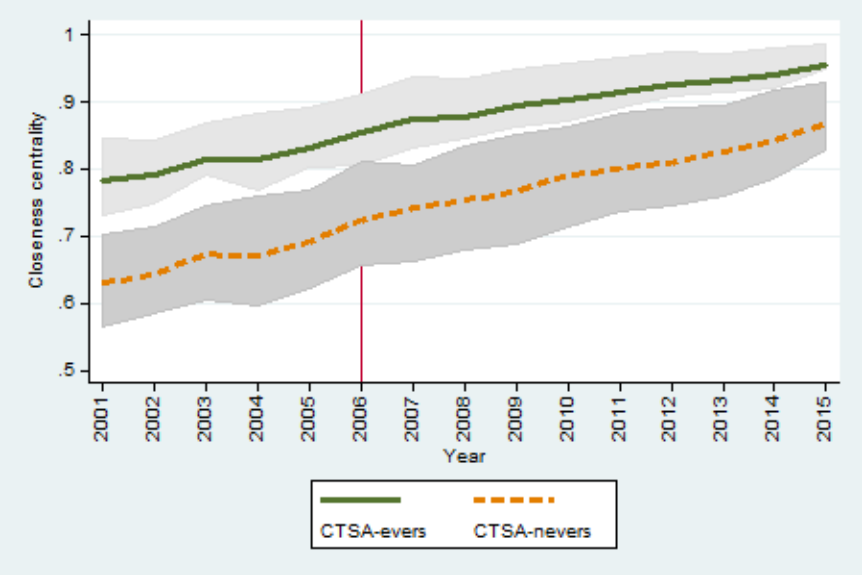


Figure 18. Closeness centrality trend of institutions in CTSA-ever group and CTSA-never group

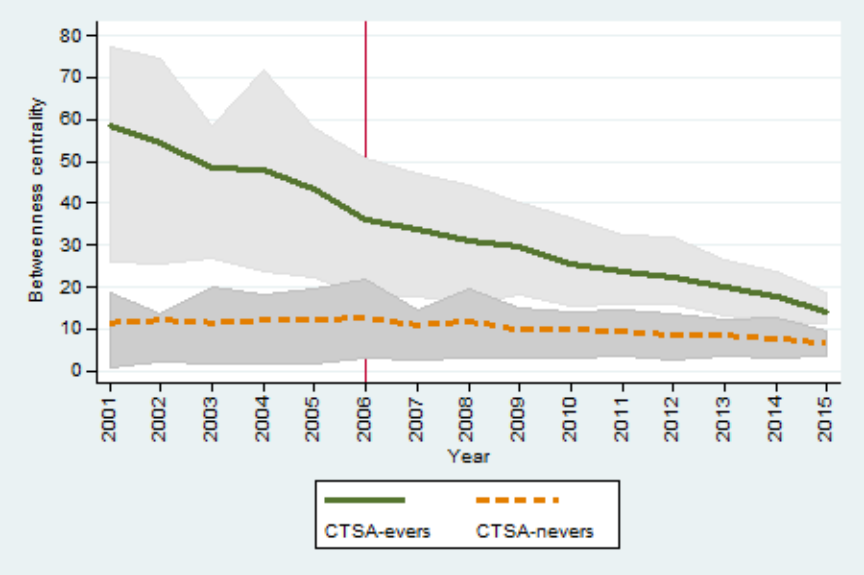


Figure 19. Betweenness centrality trend of institutions in CTSA-ever group and CTSA-never group

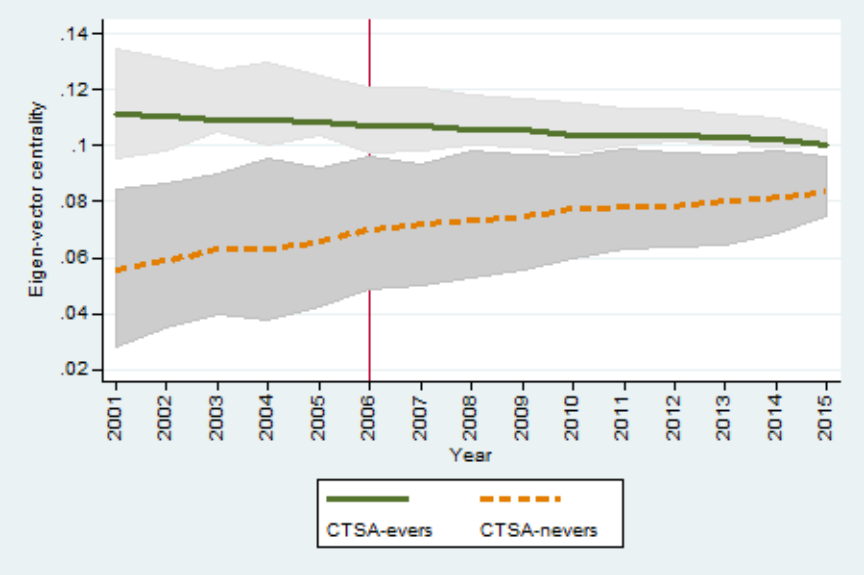


Figure 20. Eigen-vector centrality trend of institutions in CTSA-ever group and CTSA-never group

3.5.2. *Network regression*

To test whether the CTSA led to higher number of collaborators, which is Hypothesis 1, I ran ERGM models with binary dependent variable from 2009 to 2015. One of the assumptions in my analyses is that it would take three years for the impact of the CTSA award to get realized. Hence, my first year of analysis is 2009 as the first year that institutions started taking advantage of the CTSA award is 2006. Table 5 shows estimation of seven ERGM models with unweighted rate of collaboration as dependent variable. The coefficients of **CTSA** have significantly positive value for all years. This means receiving CTSA lead to more number of collaborators and hypothesis 1 holds. For instance, in 2015, CTSA increases the log odds of institutions to form a tie with other partners by 0.496[22].

If we look at the coefficient of the control variables, we can note that the effect of the R&D expenditure (**R&D**) and being private school (**Private**) have minimal or insignificant effect in making more collaborators. However, it turned out that having medical school leads to higher number of collaborating institutions. In all models, the coefficients of **Medical** were larger than that of **CTSA**. This finding indicates that the effect of having medical school has a larger impact in terms of creating more collaborators than the CTSA award.

Regarding the homophily effect, the result shows that the coefficients of two factors considered, the status of taking advantage of the CTSA award and being located in same state, have positive values with significance in 0.01 level. This means that having same status regarding the receipt of the CTSA related resources and being located in same state lead to more collaborations.

[22] The coefficients of CTSA related variables from the models that include endogenous network formation factors (e.g., preferential attachment) also showed similar values (not shown here).

Table 5. ERGM estimations with binary dependent variable between 2009 and 2015[23]

Variables	2009	2010	2011	2012	2013	2014	2015
Main effects							
CTSA	0.344***	0.296***	0.307***	0.437***	0.387***	0.485***	0.496***
Researcher	0.279***	0.269***	0.324***	0.545***	0.491***	0.516***	0.424***
R&D	0.005***	0.005***	0.005***	0.004***	0.003***	0.003***	0.004***
Private	0.009	0.072	0.056	0.219***	0.113*	0.075	0.000
Medical	0.585***	0.460***	0.465***	0.614***	0.540***	0.520***	0.477***
Homophily Effects							
CTSA	0.281***	0.369***	0.332***	0.437***	0.302***	0.401***	0.302***
State	1.714***	1.259***	1.396***	1.708***	1.606***	1.459***	1.393***
Constant	-2.608***	-2.263***	-2.254***	-2.697***	-2.161***	-2.153***	-1.679***

To test hypothesis 2 of my analysis, which is related to the impact of the CTSA award on the rate of collaborations, I ran ERGM models with weighted collaboration rate as dependent variable between 2009 and 2015. The results are shown in Table 6. From the results from models for all seven years, we could find that receiving CTSA has a significant impact on the rate of collaboration as the coefficient of **CTSA** is statistically significant for all years and it suggests that my hypothesis 2 holds. For instance, in 2009, the expected number of collaborations that the CTSA-ever institution has is $\exp(0.291) = 1.338$ times higher than that of CTSA-never institutions holding all other variables constant. In 2015, the strength of ties that a CTSA-ever

[23] Diagnostics of goodness-of-fit of models are presented in the Appendix C. It seems like the simulated networks predict the observed network reasonably well as 1) the shapes of the observed shared partners frequency are similar to those of the simulated shared partners frequency and 2) the large portion of the observed value fall within the acceptable range of the simulated values.

institution has was $\exp(0.287) = 1.332$ times stronger when compared to a CTSA-never institution, holding all other variables constant[24].

It was also interesting to note that all the control variables had statistically significant impacts. The R&D expenditure had very minimal impact whereas number of researchers, being a private school and having medical school had substantial impacts on increasing collaboration frequency. One finding that is noteworthy is the impact of medical school. We can see that the coefficient of having medical school is higher than that of receiving CTSA for all seven years. For example, in 2010 the coefficient of **Medical** is 0.685 while the coefficient of **CTSA** is 0.315. This means that having medical school leads to $\exp(0.685) = 1.862$ times more collaborations, whereas receiving the CTSA award only leads to $\exp(0.315) = 0.856$ times more collaborations when all other variables are kept constant.

The CTSA homophily effect is also significantly positive. This means that receiving the CTSA award leads to more collaborations with other CTSA recipients. That is, there is a tendency that institutions in CTSA-ever group are collaborating more with the CTSA-ever group institutions. In the previous section that explained micro-level network measure, the change of eigen-vector trend showed that CTSA-ever institutions had expanded their collaborating institutions to periphery institutions. With the result that CTSA homophily effect is always positive, we can make more advanced claim that CTSA-ever institutions have not been increasing interaction with CTSA-never institutions, but with have been increasing interaction with CTSA-ever institutions that were not in the very core of the network. This means that institutions that started to act as network hubs belong to CTSA-ever group, not CTSA-never group, even if they are in the periphery of the network. As expected, being in the same state has

[24] The coefficients of CTSA related variables from the models that include endogenous network formation factors (e.g., transitivity) also showed similar values (not shown here).

significantly large impact on creating more collaborations. In fact, the state homophily effect is larger than any other effects tested.

Table 6. ERGM estimations with weighted dependent variable between 2009 and 2015

Variables	2009	2010	2011	2012	2013	2014	2015
Main effects							
CTSA	0.291	0.315	0.252	0.312	0.297	0.290	0.305
Researcher	0.282	0.270	0.252	0.258	0.218	0.224	0.190
R&D	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Private	0.278	0.278	0.329	0.326	0.270	0.283	0.286
Medical	0.691	0.685	0.647	0.741	0.675	0.604	0.616
Homophily Effects							
CTSA	0.237	0.228	0.212	0.228	0.191	0.235	0.199
State	1.269	1.102	1.131	1.051	1.033	0.990	0.926
Constant	-1.963	-1.723	-1.625	-1.580	-1.283	-1.154	-0.909

Table 7 shows the estimations of the TERGM model. Model 1 is estimated without time-related variables and Model 2 is estimated with time-related variables. For both models, we can see that the main effect of the CTSA is positive with significance at the 0.01 level. As TERGM can only handle binary dependent variables, this means that institutions affected by the CTSA award are increasing their number of collaborators even when the temporal effects are considered. This result provides additional support to hypothesis 1 of my analysis. The impact of being a private school (**Private**) and having medical school (**Medical**) both have significant impact on institutions making more collaborators. Also, having same status regarding the receipt of the CTSA award (**CTSA** homophily effect) and being in the same state (**State** homophily effect) lead to more collaborators. All these results are consistent with the results of the yearly

ERGM estimations. The fact the constant has negative value indicates that the observed network is sparser than the model we can get from random models.

Regarding the time-related factors, both the tie memory variable and time memory are positive with significance. **Tie memory** having positive value, which is the case of my estimation, means that tie (or non-tie) at one time point lead to tie (or non-tie) on the next time point (Leifeld et al., 2018). This means the institutions are more likely to form ties with the institutions they collaborated with in the past, while not forming ties with institutions they haven't collaborated with. Therefore, positive coefficient of the **Tie memory** also means that the network is stable (Leifeld et al., 2018). The **Time** variable also had a positive value. This means that the number of connected pairs of institutions is increasing over time. This finding is consistent with the increasing trend of the whole network's density (Figure 10).

Table 7. Estimation result of the TERGM for 2009-2015

Variables	Model 1 (without time-related covariates)	Model 2 (with time-related covariates)
Main effects		
CTSA	1.296	1.068
Researcher	0.004	0.338
R&D	0.351	0.003
Control	0.075	0.111
Medical	0.476	0.421
Homophily Effects		
CTSA	1.089	1.033
State	1.459	1.127
Time-related		
Tie memory		0.677
Time		0.100
Constant	-2.895	-2.881

Note: All coefficients are significant in 0.01 level

3.6. Conclusion

In this chapter, I looked at how the CTSA award changed the landscape of biomedical research collaboration network. The first notable finding is that that biomedical research collaboration network is becoming denser. This result is in line with the findings by Powell, White, Koput, and Owen-Smith (2005) in that the biomedical research collaboration is becoming denser. One of the key goals of translational research is to break down barriers between institutions and promote interaction among them (Obeid et al., 2014). The CTSA award portfolio has programs that provide tools to promote active collaboration between institutions. One example is the Research Network System, which provides information on researchers and coauthorship activities affiliated with various institutions. (Obeid et al., 2014). With this kind of tools, CTSA tries to help researchers find their potential collaborators and acquire information on resources and methodologies (Obeid et al., 2014). It may bring the researchers affiliated with institutions with these tools to collaborate more with each other. Subsequently, this could bring more collaboration of CTSA institutions with other institutions that will lead to a denser network as a whole. The results from ERGM and TERGM confirms this claim and they indicate that receiving CTSA awards lead institutions to have more collaborators (Hypothesis 1 holds) and make more frequent interactions with the collaborators (Hypothesis 2 holds).

The second notable finding is the overall network is becoming less centralized (Hypothesis 3 hold), meaning institutions that didn't take the role of network hubs in the past are starting to take this role. The descriptive analysis of eigen-vector centrality and ERGM estimation results indicate that it is the CTSA-ever institutions that are becoming new hubs, not the institutions in the CTSA-never group. Though it was shown that having homophily feature of

receiving the CTSA award leads to more collaborations and a larger number of collaborators, the main effect of the CTSA award is stronger. Hence, CTSA receiving institutions have significantly more collaborators and collaborations than the institutions that never received the CTSA award. The reason why the CTSA-never institutions have limitation in attracting more collaboration may be explained by factors like lack of useful resources (Bozeman & Corley, 2004; Katz & Martin, 1997), not being able to provide mentorship (Katz & Martin, 1997) and not being the source of creativity (Katz & Martin, 1997). The unpopularity of CTSA-never institutions compared to CTSA-ever institution may also mean that resources, which could provide direct benefits, are more attractive than increasing the probability of being selected as CTSA recipient in the future.

Although this chapter advances our understating of the impact of the CTSA award on research collaboration, it has some limitations. Most of all, the analytical approach of this chapter is subject to endogeneity concerns. Indeed, the institutions affected by CTSA award may be the ones that have preferred characteristics in terms of collaboration (e.g., more links with other institutions) before being selected as the CTSA recipient. I tried to alleviate this concern by applying lag between the start of the CTSA award and the time when the collaboration is taking place. Also, CTSA partnering institutions, which don't necessarily have high propensity of receiving the CTSA award, are included in the treatment group which may reduce the selection bias problem. Also, I also moved beyond cross-sectional analysis and conducted temporal analysis, which could alleviate the endogeneity problem to a certain extent. However, the endogeneity concern is not fully addressed, and it allows me to discuss only the strong association of the CTSA award on the collaboration network, not the perfect causal relationship between the two.

CHAPTER 4. A MEASURE ON TRANSLATIONAL FEATURE OF PUBLICATIONS

4.1. Introduction

The ultimate goal of the CTSA award is to translate findings from the laboratory into products that benefit patients and have commercial value (Han et al., 2018; Weber, 2013). The development of new drugs, medical devices, or other therapeutics is the ultimate measure of successful translation. However, as noted in the previous chapters, it takes too long to make these products from scientific discoveries (Morris et al., 2011; Contopoulos-Ioannidis et al. 2008), for them to be an attractive measure for use in the assessment of the translation process. Patents are another potential measure that may serve as an indicator of successful translation, but the lag time for patents to issue along with uncertainty about which patents will be used commercially complicate the use of this measure. On the other hand, it takes much less time for publications describing novel research to appear, with more than 80% of publications being published within six years from the year of grant directly related to that publication (Ihli, 2016). This makes measures related to publications attractive for the evaluation of projects that support translational research. Although publications have less commercial value compared to patents and drugs, publications can be used to understand the dissemination of research and their forward citation data can be used to understand the usage of the publications across the academic discipline (Llewellyn et al., 2019). In this respect, there have been quite a lot of attempts to develop an index measuring translational feature of publication using their bibliometric information (e.g., Weber, 2013; Han et al., 2018; Fontelo & Liu, 2011; Surkis et al., 2016). However, thus far, there has been no measure that acquired broad consensus from the research community. Hence,

in this chapter, I propose a novel index that measures translational feature of publications, which has the potential to be used widely.

4.2. Background

Federal funding has supported translational research through the CTSA award for more than a decade (Weber, 2013). As these efforts have continued and represent a substantial investment, it is important to assess the effectiveness of this endeavor. Several studies have tried to assess various aspects of the impact of CTSA. For instance, some studies focussed on the research productivity and scholarly impact of publications with bibliometric data. Schneider et al. (2017), for example, used bibliometric tools to measure the impact of articles that CTSA institutions published. They found out that six CTSA institutions they analyzed experienced an increase in publication counts and forward citation counts. Another example is a work by Georgia CTSA Alliance (Llewellyn, Carter, Rollins, & Nehl, 2018), which looked at articles citing CTSA hub awards. They found that the publications supported by CTSA programs are cited more and this characteristic strengthens when a publication is from multi-institutional CTSA hubs. Other topics like clinical trial enrollment (Liu et al., 2013) and the success rate of receiving NIH grants (Knapke et al., 2015) have also been examined. These studies found that the CTSA award had a positive impact on the number of patients enrolled in clinical trials and probability of getting a NIH grant funding, respectively (Liu et al., 2013; Knapke et al., 2015).

It is important to investigate the factors mentioned above for the assessment of the CTSA program. However, the most important factor in the assessment of CTSA is whether it achieved its ultimate goal of changing the research community towards more translational research. To the

best of my knowledge, only one study has examined this issue (Weber, 2013). Using Medical Subject Headings (MeSH), Weber (2013) divided publications into three groups, which are publications on animals (group A), publications on cells (group C) and publications on humans (group H). Then, he made a triangle with A, C and H as three points. After that, he placed points that represent individual collections of publications (e.g., group of publications on biology) inside the triangle and tracked the location of points by time. If they noticed that a point of specific publication collection is moving towards the H point, they assumed that field is becoming more translational. One of the results they found is that a collection of articles that has the topic related to "Cloning, Organism" is moving towards the H-point, which means that this group of publications is becoming more translational over time. They also introduced the concept of "generations" of translation lag. They defined an article as a first-generation article if it was cited directly by group-H paper, as a second-generation article if a non-H group article that cites non-H group article got cited by group-H paper and so forth. The study by Weber (2013) provides one perspective on what has happened in the research community on translational research. However, the measurement used in the study is complex and not easy to follow and this concern should be addressed for the result to be accepted widely.

For the accurate evaluation of the CTSA program, we need to use right measure of what translational research is. Besides the study by Weber (2013), there have been few other attempts to come up with a measure of translational research. For instance, Surkis et al. (2016) made a checklist that can be used in categorizing publications into a specific stage in the five-stage process of translational research (T0 to T4). Then, they created a machine learning algorithm to categorize publications within these five stages. Due to low frequency of T1, T2, T3 and T4 articles relative to T0 articles, they combined T2 with T3 articles and T4 with T5 articles during

the classification. The authors claimed that their algorithm showed very good performance and reliability in all groups of articles. Another example using MeSH keywords for the classification of translational work is a study by Han et al. (2018). They classified a publication as "primary translational research" if it was related to field related to clinical science[25] (Han et al., 2018, p. 5). The authors also introduced a category named "secondary translational research" and classified a paper into this group if a paper was not associated with clinical study but got citations from papers that deal with clinical issues (Han et al., 2018). Using this approach, they showed that 13.4% of CTSA supported articles published in the field of behavioral and social science could be classified as translational research.

In some studies, methods not related to MeSH keywords were used for the classification of translational publications. For instance, Fontelo & Liu (2011) introduced a web application filter that can be used to retrieve articles that have potential clinical applications. They created this filter by manually reviewing words and phrases that appear in articles published in clinical and translational science journals. A study by Grant, Cottrell, Cluzeau, and Fawcett (2000) looked at the citation link between publications and clinical guidelines. They assumed that publications cited by clinical guidelines have an impact on the field of health. Based on the type of journal in which an article was published, they classified cited publication as 1) clinical observation, 2) basic, 3) clinical mix, or 4) clinical investigation. Their result showed that publication in "basic" category is not cited much (8%) by clinical guidelines.

Looking at previous studies, we can easily notice that the research assessment field has not standardized on a single indicator to measure the translational feature of research articles and the attempts to develop and improve measures are still ongoing (Blümel, 2017; Surkis et al.,

[25] Few examples include clinical study, clinical trial, phase 1 clinical trial, phase 2 clinical trial, phase 3 clinical trial, phase 4 clinical trial, controlled clinical trial, practice guideline, observational study and randomized controlled trial

2016). Hence, when a researcher wants to conduct research on the change of translational feature, he or she should select an existing measure or develop a novel measure that fits the context of the planned research well. I made the latter choice in my study and described this measure in the sections that follow.

4.3. A new measure: the TS score

To measure how translational a publication is, I am proposing a new index that is related to the share of forward citations from clinical science a non-clinical article receives. This measure aligns with the method of using disciplines of journals of the citing article to characterize features of the cited article (Qin et al., 1997; Grant et al., 2000). Two big steps are required to calculate the new measure. As the first step, I categorized journals into four disciplines, which are clinical science (Clinical), non-clinical science (Non-Clinical), multidisciplinary (Multi) and non-science and engineering (Non-S&E). This categorization is done by using Web of Science Category (WOSC) and NSF classification of fields of study (National Science Foundation, 2018). In Web of Science, each journal is assigned up to five WOSCs based on the topics covered by the journal. Each WOSC can be matched with fields in the NSF classification of fields of study using the name similarity. After the name matching, I classify the Web of Science category of journals into clinical science and non-clinical science. Specifically, when a field was listed in science and engineering section (e.g., biomedical engineering) I classified it as non-clinical science and, when the field was listed in health section (e.g., pediatrics), I classified it as clinical science. As a journal can have up to five WOSCs, a journal could have a mix of non-clinical science WOSCs and clinical science WOSCs. In my

study, I classified a journal as clinical journal if the journal has only clinical WOSCs. The fields that were not listed under science and engineering or health were classified into Multidisciplinary science or Other Science (e.g., social science)[26]. Figure 21 illustrates the concept of the first step.

Figure 21. Classifying publications into the ones with and without translational feature based on journal discipline of forward citing publications

As the second step, the Translatedness score (TS score) is calculated using the simple equation shown below.

$$TS\ score = \frac{Number\ of\ forward\ citations\ received\ from\ clinical\ science}{Total\ forward\ citations\ received} \quad (3)$$

Using the portion of cross-disciplinary citation builds on the approach that Luke et al. (2015) adopted. In their study on the cross-disciplinary collaboration of an individual CTSA receiving institution, Luke et al. (2015) introduced the cross-disciplinary density ratio, a measure calculated by dividing the density of cross-disciplinary collaboration by the density of within-disciplinary ties. However, the focus of their study was not on articles but on the composition of research teams.

The intuition behind the TS scores is that if an article published in a basic science journal has translational feature, it is likely that it receives a large portion of forward citations from

[26] Details on the classification of Web of Science Category into four disciplines are shown in the Appendix D.

journals published in the clinical sciences. Therefore, there will be a baseline TS score to judge whether an article has translational feature or not. This approach would provide me with the number of translational publications each institution produces every year and I could track the change of this count variable to evaluate how much an institution is impacted by the CTSA award.

4.4. Verification tests on the measure

When a new measure is developed, we need to assess how credible the measure is. In this regard, I've conducted some tests that can give support to the reliability and validity of the measure.

4.4.1. *Reliability tests*

What is essential for a quantitative measure to have high reliability is the stability over time (Golafshani, 2003). A measure's time stability is related to similarity of measurements within a given period of time, which could ensure repeatability and replicability of a measure (Golafshani, 2003). In this regard, we need to check if the TS score has stable value over time in order to claim that it is a reliable measure. For each publication, enough number of forward citations should be accumulated for the share of forward citations from clinical science to reach a stable status[27]. As the forward citation counts will increase over time, the TS score of an article

[27] For instance, if a publication received first two citations from non-clinical field, its TS score will be 0. If its third forward citation is from clinical side, its TS score will rise from 0 to 0.33. On the other hand, if the forward citations count of a publication is large enough the TS value would not change much with one additional forward citation from rare discipline. For instance, let's assume a situation where an article received 100 forward citations and all of them were from non-clinical publications. If its 101st forward citation comes from clinical side, its TS score will only increase from 0 to 0.01.

will stabilize only after certain years has passed since the publications of the article. This means that forward citation counts and time since publication, which is correlated with each other, are two factors to consider in determining the stabilization of the TS score. In this respect, I took the following steps to find out when the TS score gets stabilized[28].

1. Check the TS score of each publication by year and figure out how many years are needed for the TS score to stabilize, which I call *stabilizing year*.
2. Extract groups of papers with a certain number of forward citations (i.e., 1 through 10) and calculate the TS scores of each publications-group. This group of TS scores will be named as *TS scores at starting year*.
3. Calculate the change of TS scores of each publication-group by year (i.e., *N* years) for the years after *stabilizing year* (calculated in step 1), which I name *TS scores after N-year*.
4. Check the correlation of *TS scores at starting year* (acquired in step 2) and *TS scores after N year* (acquired in step 3). If the correlation is larger than 0.7 for all *TS scores at starting year - TS scores after N-year* pairs, we could assert that the TS score was stable from the starting year.

To perform step 1 through 4, I used a sample of 286 non-clinical publications that were published in 2003 and acknowledged NIH grants with a PI from Emory University. The cumulative counts of total forward citations and forward citations from the clinical science journals that occurred between 2004 and 2015 were used to calculate the TS score. Figure 22 shows the change of TS score by year for some publications. These are some examples of publications that received forward citations for almost all years (11 or more years) during the

[28] Ideally, it's better to track the TS score by the change of the forward citation counts. However, I was only able to get end-of-year values of forward citation counts and the values were not the number that increment by one. Hence, I had to use the method presented below.

period of interest[29]. There are various forms of change regarding the TS score. Some publications have similar TS score from the beginning till the end (e.g., ID 272, ID 415), some publications start from very low TS score value but increase as time passes (e.g., ID 36, ID 195) and some publication starts with high TS score in the early years and decrease over time (e.g., ID 13, ID 267). TS score of some papers that fluctuate in the early years (e.g., ID 24, ID 269). However, we can note that, in almost all cases, the value of TS scores doesn't change much after 2007, which is marked as the vertical line in each graph. Having these results, we can claim that the TS score of a publication reaches its stable value four years after a paper is published. It means that we need to wait around four years to classify a basic science publication into one with high translational feature or the one with small or no translational feature.

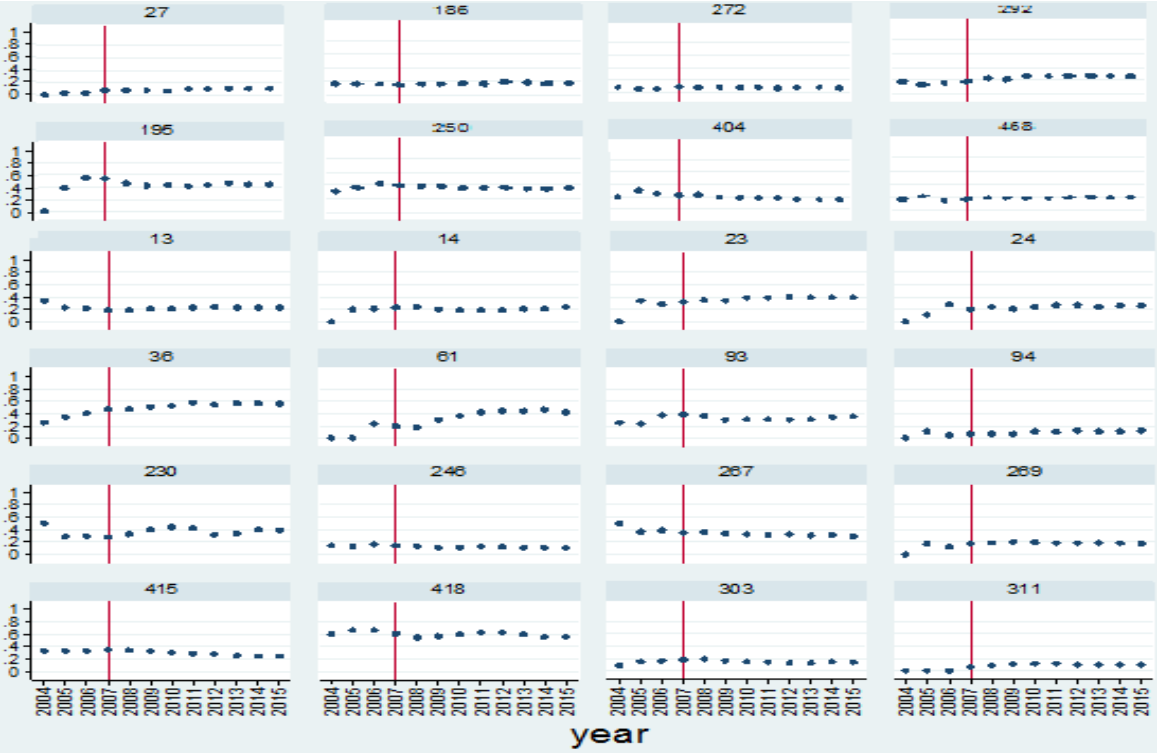


Figure 22. Change in TS score by year for some non-clinical publications

[29] More examples can be found in the Appendix E.

Table 8 shows the result of step 2 through step 4 of my reliability test. If we look at the case of publications with only one forward citation, which is a sample with 63 cases, the correlation of the TS scores at the starting point and four years and five years after the starting point is 0.730 and 0.721, respectively. However, the correlation between starting point's TS scores and six-year after starting point is 0.634, which is below 0.7 threshold. Furthermore, the correlation value stays below 0.7 if we increase the post-year to 7 and 8. Hence, having only one forward citation doesn't give us a stable TS score, as expected. This is also the case for the publications with two citations. The correlations are smaller than 0.7 threshold for all starting point and post starting point pairs. For citations with three citations, the correlation values are all larger than 0.7, which seems to be a coincidence. For the groups of papers that start with four citations, the correlation values are smaller than 0.7 in most pairs (all pairs except starting point and after five years pair). The tendency of having correlation value smaller than 0.7 stops as the forward citation count at start point increases to 5. For the group of publications that received 5 or more forward citations, the correlation value is larger than 0.7 for all pairs. Therefore, we could claim that the publications with five or more forward citations have stable TS score values.

Table 8. Correlation between TS score of the starting point and TS score after four or more years

Forward citation counts at starting year	**Correlation between TS scores at the start and corresponding years**					**Sample size**
	After 4 years	**After 5 years**	**After 6 years**	**After 7 years**	**After 8 years**	
1	0.730	0.721	0.634	0.589	0.550	63
2	0.661	0.504	0.533	0.512	0.462	62
3	0.794	0.799	0.770	0.732	0.727	71
4	0.690	0.707	0.637	0.645	0.622	81
5	0.786	0.745	0.789	0.793	0.808	69
6	0.836	0.792	0.785	0.770	0.743	59
7	0.827	0.807	0.800	0.810	0.812	74
8	0.904	0.891	0.879	0.874	0.849	50
9	0.878	0.865	0.853	0.833	0.831	64
10	0.915	0.923	0.918	0.931	0.937	55

4.4.2. *Validity tests*

A. *Calculating the TS scores with another approach: Publication discipline classified based on the affiliation of the first authors*

Validity of an indicator is related to the question that determines whether an indicator is measuring what it intended to measure (Golafshani, 2003). The validity in quantitative research like mine is associated with truth, actuality and accuracy (Golafshani, 2003). The TS score I am proposing is based on the classification of disciplines of journals. If the new TS scores calculated through another method have similar values (high correlation with my original TS score), the validity of the measure will increase. There have been some other measures that examine citations outside the category an article was published in, mainly for assessing the multidisciplinarity of articles (e.g., Morillo, Bordons, & Gómez, 2001; Ortega & Antell, 2006; Porter & Chubin, 1985). One of the approaches that these studies used is taking the first author's affiliation to categorize papers into different disciplines (Ortega & Antell, 2006). This assumption is based on the accepted practice in academia, especially in life science, that first authors of publications are assumed to have the largest contribution on the paper (Carpenter, Cone, & Sarli, 2014, p. 1162). Drawing on these articles, I recalculated TS score using the discipline classification of the first author's affiliation to help assess the validity of my measure. For this analysis, I classified forward citing publications with the first author affiliated with clinical science department as clinical science publications and classify articles as non-clinical publications otherwise. Figure 23 depicts the concept of classifying publications based on the first author's affiliation.

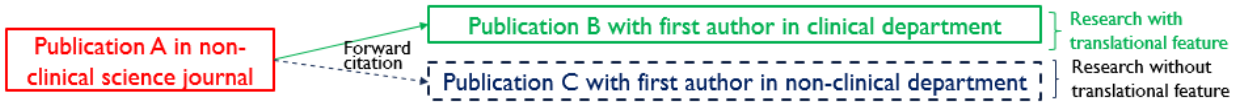


Figure 23. Classifying publications into the ones with and without translational feature based on first author's affiliation

The steps that I took to conduct this validity test are explained below:

1. Select a sample of publication for the analysis. My selection is journal articles that were published in non-clinical science journal in 2003 that acknowledge any NIH grant that have PI from Emory University with forward citation counts between 11 and 42. This gave me a sample of 114 articles.
2. Gather the bibliometric information of the articles that cited the publications in my sample. This gave me a total of 2,089 forward citing articles.
3. Classify disciplines of forward citing publication based on the full address of the affiliation of the first authors. There were around 2,000 unique affiliations, which I classified manually[30].
4. For each article, calculate the share of forward citing publications that have first author's affiliation classified as clinical department.

[30] There were some issues to consider when classifying publications into clinical papers and non-clinical papers, which led me to do the classification manually. First, the department information didn't exist in some listings (e.g., Univ St Andrews, St Andrews KY 16 9ST, Fife, Scotland; RAND Corp, Santa Monica, CA 90401 USA). In these cases, I couldn't make decision on the discipline of the author's affiliation. Second, the orders that university name, department or school name and city are listed were different between articles. Though the department information was listed second in most cases, it was listed first in some cases. Third, the details of the affiliation were different across publications. For instance, some publications provided very detailed information (e.g., Sch Med & Dent, Dept Biostat & Computat Bio) whereas some publications only provided information on school level (e.g., Sch Life Sci & Technol). To be conservative on classifying a publication into clinical paper, I classified the publications into clinical article only if there were clear and enough information of the discipline. For instance, I didn't classify a publication as clinical if the most detailed information of the first author's affiliation is school of medicine. This is because, in many universities, school of medicine is composed of departments conducting basic science work (e.g., department of microbiology). In this vein, I only classified a publication as a clinical paper only if sub-division of the school is closely related to the clinical field of research (e.g., Department of Pediatrics). This approach would give me lower values for the TS scores.

Table 9 compares summary statistics for two different approaches, one based on the journal classification of forward citing articles and the other based on the affiliations of first authors in forward citing articles. We can see that the value of the TS score based on first author's affiliation (first row) is larger than the TS score calculated based on journal classification (second row). A total of 81 articles out of 144 articles (71.1%) in my sample had larger TS scores when the first author's affiliation-based approach was used. Figure 24 shows the distribution of TS scores calculated using two different approaches. We can see that the TS scores calculated using the first author's affiliation has larger mean value and are more dispersed, as expected. Most importantly, the correlation value between two TS scores is 0.748, which is larger than 0.7 threshold. Hence, we can claim that TS score based on journal classification has gained validity in some sense.

Table 9. Correlation between TS scores calculated with different methods

Approach of classifying discipline of forward citing article	**Mean**	**Std Dev**	**Min**	**Max**	**Correlation with TS score based on journal classification**
Based on journal classification	0.092	0.119	0	0.553	1.000
Based on first author affiliation	0.172	0.164	0	0.684	0.748

Note: The number of observations is 114 for both groups

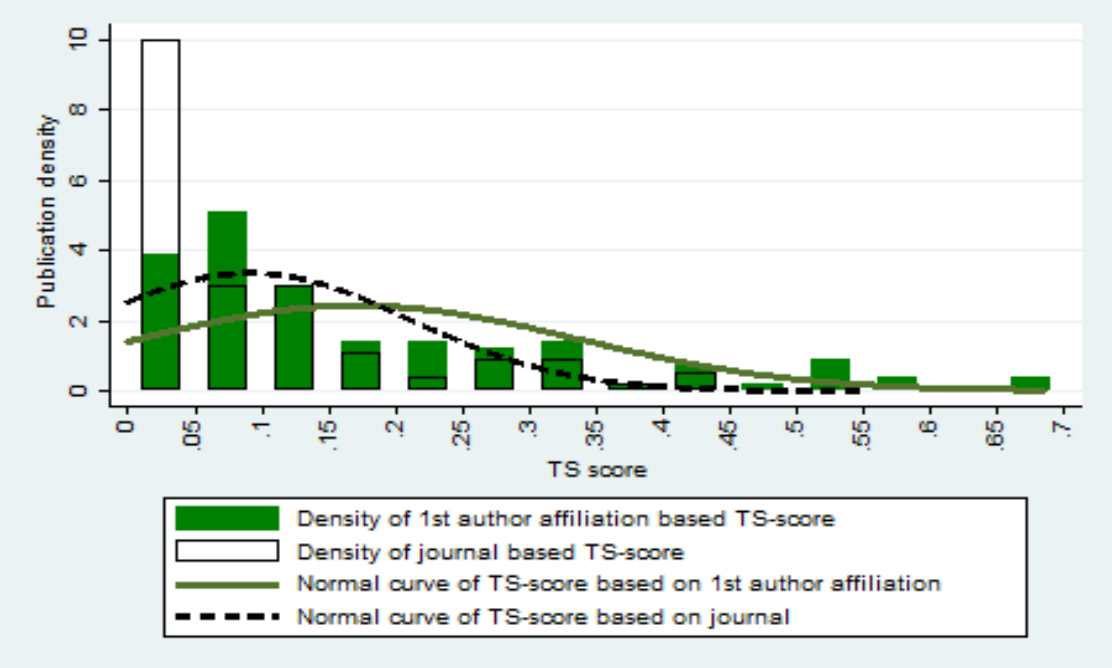


Figure 24. Distribution of TS scores calculated using different approaches

B. Checking the TS scores of publications that result in patents

If a publication is cited by one or more patents, it suggests that the publication has a translational feature, or is at least relevant to industrial use (Meyer, 2000). A direct knowledge-transfer from basic research to technological outcome could be reflected in patent's citation of basic science publication (Huang, Yang, & Chen, 2015; Meyer, 2000; Sung, Wang, Huang, & Chen, 2015). US patents list nonpatent references (NPR) on their front pages and the citation link between publications and patents can be identified using the NPR list (Meyer, 2000). Drawing on this evidence of patenting and translation of research, I aim to test the validity of my TS score using publication-patent citation link, using both a prospective approach and a retrospective approach.

Prospective approach of publication – patent citation analysis

As the first step to test the validity of the TS score using prospective approach of publication-patent citation analysis, I selected a group of basic science publications. My selection was, as in reliability tests, the set of 286 non-clinical articles published in 2003 that received support from NIH and had a PI affiliated with Emory University. I searched the DOI of these 286 articles in LENS.ORG (http://www.lens.org)[31]. The Lens returned information of 217 of them giving me the retrieve rate of 75.9%. Not all publication information is retrieved as the Lens database doesn't cover all scholarly databases, but only covers three datasets (PubMed, Crossref and Microsoft Academics)[32].

Table 10 shows the summary of the statistics for publication groups with and without patent citation. Among the total sample of 217 publications, 144 publications (66.4%) didn't receive any patent citation and 73 publications (33.6%) had at least one patent forward citations. We can see that the average TS score of the publications without any patent citation (0.111) is smaller than the average TS score of publications that result in at least one patent (0.156). The t-test also showed that the mean value of the TS score of the former group is smaller than the mean TS score of the latter group in 0.05 significance level (Pr value: 0.013). This result provides additional support that the TS score is a valid measure of the translational feature of publications.

[31] LENS.ORG is an open public website managed by Cambia, an independent non-profit organization, that provides linkages between scholarly works, patents and biological sequences (LENS.ORG, n.d.). Its patent database covers patent datasets from the USPTO, European Patent Office, WIPO and IP Australia and its scholarly dataset includes PubMed, Crossref and Microsoft Academics (LENS.ORG, n.d.). With the collaboration with NIH Pubmed and Crossref teams, the Lens linked publications' Digital Object Identifier (DOI) with NPL in their patent database. Hence, using DOI of publications, we could check if the publications were cited by patents or not.

[32] Using Pubmed ID instead of DOI for the search may give me higher retrieve rate. However, it would be hard to track the forward citations links of publications with only Pubmed ID but without DOI.

Table 10. TS score comparison between publications resulting and not resulting in patents

Group of publications	Obs	Mean	Std Dev	Min	Max
All publications	217	0.126	0.139	0	0.611
Publications without patent citation	144	0.111	0.126	0	0.610
Publications with patent citations	73	0.156	0.158	0	0.611

Retrospective approach of publication – patent citation analysis

For a test on the validity of TS-scores using a retrospective approach of publication-patent citation analysis, I started with group of patents, rather a group of publications. To complete this validity test, I used the following approach:

1. Using patent ID - NIH project ID link CSV file downloaded from NIH RePORTER, acquire the list of patents issued during 2001 -2015 that acknowledge any NIH project that have Emory University affiliated PI (List 1).
2. Using "otherreferences" file downloaded from PatentsView (2018), extract bibliometric information (e.g., author, title, journal names) of "Non-patent citation" of List 1 patents (List 2).
3. Using bibliometric information obtained from step 2, manually search for non-patent citations using Google Scholar and Web of Science to get their DOIs (List 3).
4. Extract publications in List 3 that are from basic science journals (List 4).
5. Get forward citing publication list of DOIs in List 4 (List 5).
6. Calculate TS scores of DOIs in List 4 using the Web of Science Category classification of List 5 publications.

The result from step 1 gave me a sample of 177 patents (List 1). This group of patents had a total of 186 non-patent citations (List 2), of which 88 were scholarly publications (List 3). Among 88 scholarly publications, 68 of them were published in basic science journal based on my classification rule using the Web of Science category and NSF Classification of fields of study (List 4). 68 basic science publications received 29,503 forward citations (List 5), which were divided into clinical publications and non-clinical publications based on the Web of Science category and NIH classification of fields of study. Table 11 shows the summary of all basic science publications (sample size 68) and group of publications that received more than five forward citations (sample size of 66). This validity test benefits from focusing on publications that have been cited in patents but lacks a natural comparison group. A matching approach could potentially be used to develop a comparison group but that was beyond the scope of this dissertation. Hence, I compared the mean value of this group with the average TS score of the whole group of publications[33] that will be used in my analysis in chapter 5. The mean values (0.135 and 0.140) were larger than the average value of the whole group of publications, which was 0.117. Hence, this finding could be used as additional evidence of the validity of the TS score I am proposing.

Table 11. Summary of TS scores of Emory University's non-clinical publications that result in patents

Group	Obs	Mean	Std Dev	Min	Max
All publications	68	0.135	0.163	0	0.624
Publications with more than 5 forward citations	66	0.140	0.163	0	0.624

[33] All articles published during 2001 - 2015 with support from NIH grants that have PI affiliated with from 115 Carnegie R1 universities

C. Several examples of the TS scores of publications that result in drug

One very clear way to determine whether a publication has translational feature or not is checking if the publication leads to the development of a drug. Scholars have focused on this characteristic of publications and tried to find articles that result in drugs. For instance, Sampat and Pincus (2015) used a machine learning technique to match publication, patents and drugs. They did this to figure out how publications from Academic Medical Centers (AMC) are used and found that about half of FDA approved new molecular entities (NME) between 2000 and 2009 are associated with publications from NIH-funded AMC projects. Another study by Li et al. (2017) linked publications with patent associated with approved drugs. They found that around 5% of NIH grants result in publications that are cited by patents related to marketed drugs.

If the TS scores of basic science publications on abovementioned lists turn out to be high, it will give additional support to the validity of the measure. I wasn't able to get a full list of publications that result in drugs as the studies looking at publication-drug link mentioned above did not provide enough details for the exact replications. However, I was able to get some examples of publications that end up in drugs from the supplementary material of Sampat and Pincus (2015). In the supplementary material, they provided a list of 15 publication-patent-drug link examples. With the information, I calculated the TS scores of all four publications on the list that were published in non-clinical science journals based on my classification rules. Table 12 shows the value of TS scores of these publications. We can see that, in general, the TS scores have high values with the largest value of 0.552. The average TS score of these publications is 0.393. This value is much higher than the mean value of total population of papers that will be used in my analysis in chapter 5, which is 0.117. Though the number of publications analyzed is

very small, this finding gives additional support of the validity of the TS score I have developed and proposed.

Table 12. TS score of non-clinical science publications that result in drugs

Brand	Drug	Cited Article	TS score
Tasigna	Nilotinib Hydrochloride Monohydrate	Bhat et al. (1997)	0.552
Entereg	Alvimopan	Bagnol, Mansour, Akil and Watson (1997)	0.528
Inomax	Nitric Oxide	Garg and Hassid (1990)	0.345
Velcade	Bortezomib	Rock et al. (1994)	0.149

Note: Table recreated from a table in supplementary appendix of Sampat and Pincus (2015)

4.5. Conclusion

In this chapter, I introduced the TS score, a novel measure to assess the translatedness of a non-clinical publication. The measure is calculated using the share of clinical forward citations among all forward citations a non-clinical article receives. This is based on the assumption that if a non-clinical article is cited by an article in clinical science journal, a knowledge flow from basic science to clinical science is taking place. The measure borrows concepts from previous studies using journal discipline of forward citing articles (Qin et al., 1997), distinguishing basic research and clinical science (Grant et al., 2000; Narin, Pinski, & Gee, 1976), and using cross-disciplinary density ratio (Luke et al., 2015; Porter & Chubin, 1985).

Several considerations make the TS score both novel and strong. First, the measure is relatively simple and easy to understand. Potential users only need to know the concept of forward citation and the classification of journal disciplines based on fields of study. Second, the measure can be used easily. Previous measures on translational feature of publications used machine learning technique (e.g., Surkis et al., 2016), complex application of MeSH keywords (Weber, 2013; Han et al., 2018) or filters created by manual review of articles (Fontelo & Liu, 2011). All these approaches may have good precision, but it would be difficult for the common user to understand and, hence, require supplementary training. On the other hand, the TS score is straightforward in the sense that it uses very few concepts in the calculation. Also, the calculation process could be automated easily as long as there is consensus on the classification rule of journals into clinical group and non-clinical group. Third, my preliminary assessments suggest the measure is both reliable and valid. My reliability tests showed that if a publication received five or more forward citations, the TS score of the publication reaches a stable value. Results from the validity tests that compared my TS score with the TS score calculated using the first author's affiliation of forward citing publication and that checked the TS score of publications that result in patents or drugs also showed positive results. All in all, the TS score I am proposing has many strengths and I expect that this measure has the potential to be used widely for the evaluation of the translation process of biomedical research.

CHAPTER 5. IMPACT OF NIH'S CTSA AWARD ON TRANSLATIONAL RESEARCH PERFORMANCE OF INSTITUTIONS

5.1. Introduction

Creating research outcomes that have translational nature is the ultimate goal of NIH's support for translational research (Han et al., 2018; Weber, 2013). In this chapter I aim to test the hypothesis whether the CTSA award has led to a change in the number of publications that have translational feature. To classify publications into those with and without translational characteristic, the TS score that is proposed in the previous chapter will be used. A non-clinical science publication will be considered as having translational feature if its TS score is above certain threshold. To compare the translational research performance of Carnegie Highest Research Universities, all NIH supported basic science journal papers published by these institutions between 2001 and 2015 are analyzed. Difference-in-difference methodology with institution matching will be used to test the difference between the treatment group, which are institutions that received CTSA awards, and the comparison group, which are institutions that did not receive or participate in CTSA awards.

It may be the case that a variety of factors mediate the relationship between the CTSA award and the translational publication counts of institutions. Hence, in addition to assessing the relationship between the CTSA award and the production of translational outcomes, I aim to examine the process through which this relationship is formed. NIH emphasizes inter-organizational collaboration and multidisciplinary research as very important factors for the successful execution of the CTSA program (National Institutes of Health, 2017b; Obeid et al.,

2014). Hence, I plan to assess if these two factors affect the relationship between the CTSA award and the increase of translational research outcomes of institutions. For the test of mediation, a combination of methods proposed by Barron and Kenny (1986) and Preacher and Haye (2004) were used. I expect that the study will provide insight into the effects of the CTSA program and provide implications on the government policy related to translational research.

5.2. Theoretical background, previous studies and hypotheses

5.2.1. *The direct relationship between the CTSA award and translational outcome*

Research outcomes should be benchmarked to supporting program's intended aims (Corredoira, Goldfarb, & Shi, 2018) and the ultimate goal of the CTSA award is to produce research outcomes that link the outcomes from basic science to clinical use (Han et al., 2018). My first research question for this chapter addresses this fundamental goal of the CTSA program. I want to find out whether receiving the CTSA award led institutions to create more research outcomes that transfer the research outcome from the basic side to the clinical side. For this assessment, I focus on the creation of journal articles that are translational in nature[34]. To change the feature of the whole institution at large towards more creation of translational research outcomes, we need to understand what are the factors that could lead to the change. Hence, I try to identify and discuss the factors that could take the role of changing the scientist's research behavior, which would subsequently lead to the change of the overall nature of the research community.

[34] The reasons why I chose publications, not other measure like patents, for the assessment are provided in chapter 4.

The theory that connects many of the factors related to potential changes in researcher's behavior is the incentive theory. Incentive theory maintains that people would be willing to adjust their behaviors that would bring them more benefits (Balconi et al., 2004). If a scientist thinks that certain type of research would not provide them any benefit, he or she will not participate, even with intervention from the external forces (Brint, 2005). The following paragraphs list and explain the factors that may have effect on incentivizing scientist to conduct more translational research.

Meeting environmental demand. Scientists try to adapt to environmental demands (Bessant et al., 2001). As members of the society, scientists should continuously pursue the roles that the society or the affiliated institutions require and try to meet their needs (Wowk et al., 2017). Scientists take the role of "supplying" knowledge and information to the society and this should correspond with the "demand" of the society (Sarewitz and Pielke, 2007). In that regard, in the era of more emphasis on the commercial use of scientific outcomes (Brint, 2005), referred as "Model 2 knowledge production" in the study by Gibbons, Nowotny, Schwartzman, Scott and Trow (1994), scientists may change their research behavior towards more production of research outcomes that have practical uses. To achieve this goal, scientists may get more involved in translational research rather than being engaged in the traditional field they have been working on. This is also related to the self-interest theory in the sense that people would have a preference in the research type that provides them enough resources for research. As noted in previous chapters, the support for translational research continues to grow in terms of budget size, which leads to an increase of related resources. This trend could make translational research become more popular than before and attract more scientists.

Financial benefit. There are various types of benefit that scientists will react to and one of them is the financial benefit (Van Rijnsoever & Hessels, 2011). The question of whether financial benefit, represented by external funding, can lead scientists to change their research behavior to the form that they are not used to is very important in science policy. The answer to this question is crucial as various institutions and funding agencies have been trying to lead the scientists to conduct scientific research towards the direction they want. If scientists only stick to the research patterns they are accustomed to and are reluctant to change their research behavior even with financial incentive, making programs and providing grants that aim to make this change would have limited effects. The CTSA program, like many other NIH programs, goes through reassessment process every five years (Kaiser, 2017). Hence, the scientists involved in the CTSA program would tend to follow the requirements of the NIH for the renewal of the program. The most comprehensive requirement that the NIH provides to the CTSA receiving institutions is to conduct more translational work (Han et al., 2018). As a result, scientists, especially those in the leadership of individual institutions' CTSA program, would be conducting translational research while their institutions are receiving the CTSA award.

Duration of research. The third type of benefit related to researcher's incentive is the expectation on the duration of the research. Academic researchers would prefer to be engaged in the field that they can participate throughout their career, not the ones that are not expected to last long. Therefore, if they consider the movement towards translational research is temporary, they would stick to their original research field (Luke et al., 2015). However, several factors suggest that the support for translational research will last for an extended period. First, the emergence of support for translational research did not happen suddenly. It was a result of the long-lasting recognition of the problem of the "valley of death" (The President's Council of

Advisor on Science and Technology, 2012). In this respect, public authorized institutions like NIH, IOM and even the Congress recognize the importance of translational research (Rubio et al., 2010; Han et al., 2018). Also, it has been more than two decades since it gained attention from the research community as a solution to the decreasing productivity of biomedical research (Lander & Atkinson-Grosjean, 2011). Given this circumstance, support for translational research is expected to last for a long time and scientists would also be aware of the fact.

Recognition. One of the factors related to intrinsic motive of scientists that result in the change of scientist's research behavior is the recognition. Recognition of a scientist among his or her colleagues is important as it is closely related to promotions such as receiving tenure and full professorship (Butler, 2008; Van Rijnsoever & Hessels, 2011). Therefore, academic researchers would not be willing to participate if certain type of research if it does not have a positive effect on their recognition (Van Rijnsoever & Hessels, 2011). As translational research is relatively new to the entire biomedical field, questions remain about whether it will have positive or adverse impact on the recognition. As there is uncertainty, we can expect that scientists may not actively engage in translational research.

Table 13. The expected direction of the impact of factors on the change of scientist's behavior towards translational research

Factor	Direction of the change towards translational research
Meeting environmental demand	Positive
Financial benefit	Positive
Duration of research	Positive
Recognition	Not clear

Table 13 shows the summary on the expected direction of impact of the factors related to the change of scientists' behavior towards translational research. The counts of publications with

translational feature will increase, decrease or don't change depending on the relative influence of each factor mentioned in Table 13. Though the direction of the impact regarding recognition factor is unclear, the remaining factors are expected to have positive impacts towards more translational work. Hence, I am proposing following hypothesis.

Hypothesis 1. **Academic institutions receiving CTSA award will produce more translational publications than the ones not receiving CTSA award.**

5.2.2. *The indirect relationship between the CTSA award and translational outcome*

One issue with Hypothesis 1 is that the relationship mentioned in the hypothesis cannot stand alone. Many other factors could affect the relationship between CTSA award and research outcomes. In addition to traditional covariates, such as the research capacity of institutions or the existence of medical schools, important point to consider when analyzing the relationship between the CTSA award and translational publication count is to consider the mediating role of the factors related to inter-organizational collaboration and multidisciplinarity of institutions.

Inter-organizational collaboration as potential mediator[35]

Universities possess imperfect information and this shortage leads to an inefficient and unorganized form of the institution, which makes them seek partners. Powell, Koput, & Smith-Doerr (1996) urged that network will bring benefits to all stakeholders in the field of

[35] More details about the characteristics, impact and trend of inter-organizational collaboration related to biomedical research, especially on the CTSA award, can be found throughout chapter 3.

biomedicine. They urged that the network will bring more access to knowledge for all parties in the network and help them perform better. It is reasonable to expect that these resources will help scientists enhance their research productivity (Katz & Martin, 1997; Gaughan et al., 2018). The elements of collaboration that have positive effect on research productivity include division of labor, research capacity or complementary skills, intellectual stimulus, new skills learned from the collaborators, access to equipment (Lee & Bozeman, 2005; Owen-Smith & Powell, 2003).

There has also been an emphasis on research collaboration by NIH and one of the major goals of translational research is to foster collaboration between institutions (Obeid et al., 2014). This is based on the wide-spread belief that collaboration will bring many benefits such as access to resources and mentorship (Bozeman & Corley, 2004; Katz & Martin, 1997; Katz et al., 2004). Many studies also showed empirical evidence that groups with more ties with other groups have higher performance than the ones with fewer ties (Katz et al., 2004). Despite this evidence, we should remember that increasing collaboration is not the ultimate goal of the CTSA program. Rather, it's just part of the process of creating more research outcomes that have translational feature. In this regard I am proposing following hypothesis with the depiction of the hypothesis presented in Figure 25.

Hypothesis 2. **Intensity of inter-organizational collaboration of institutions mediates the relationship between receiving CTSA and production of translational outcomes.**

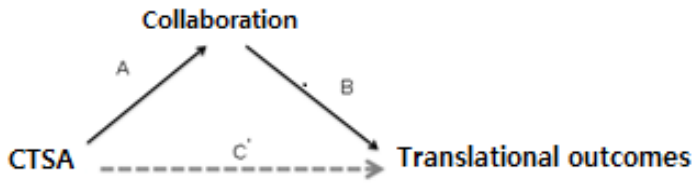


Figure 25. The depiction of Hypothesis 2

Multidisciplinarity of research as potential mediator

Emphasis on multidisciplinary research has attracted the interest of the research community for the last few decades (Qin et al., 1997; Wang et al., 2017). This can be explained with wide-spread belief that assembly of knowledge from various disciplines would lead to solutions to complex problems (Baumwol, Mortimer, Huerta, Norman, & Buchan, 2011; Van Rijnsoever & Hessels, 2011). Scientists felt the need for more combination of knowledge from diverse disciplines to fully answer important questions that traditional approach on science cannot provide (Aboelela et al., 2007). Studies on multidisciplinary research started as early as the 1950s by sociologists, psychologists and science historians (Qin et al., 1997) and have found that the major benefits of multidisciplinary research include its ability to spark creativity and make scientific breakthroughs (Rafols, Leydesdorff, O'Hare, Nightingale, & Stirling, 2012; Heinze, Shapira, Rogers, & Senker, 2009; Hemlin, Allwood, & Martin, 2004; Stevenson et al, 2013), support innovation (Gibbon et al., 1994) and help address societal problems (Lowe & Phillipson, 2006; Rafols et al., 2012; Wowk et al., 2017).

Thanks to positive perception, various parties provide support for multidisciplinary research. Government is the key player in supporting multidisciplinary research. For example, NSF considers multidisciplinary research as a major tool in tackling complex science issues and solve societal problems (Wagner et al., 2011). In the similar vein, NIH identified multidisciplinary research as their priority in their Roadmap and considered it as a crucial factor in producing new knowledge (Aboelega et al., 2007). Various funding instruments for multidisciplinary research are being applied at the university level (Van Rijnsoever & Hessels,

2011). In particular, both NSF and NIH supported universities in making large multidisciplinary research groups and building research centers for them[36] (Brint, 2005).

In line with the support from the government, there have been movements by the universities on their own. As a part of the academic revolution, universities are making programs fostering multidisciplinary collaboration and trying to become new foci of this new concept of research (Brint, 2005). During the period of increasing ties with the commercial actors (Brint, 2005), universities are taking a new approach of introducing programs that promote multidisciplinary collaboration such as making research umbrella groups and setting new initiatives. Brint (2005) mentioned that universities are also creating graduate programs dealing with multidisciplinary research and showed that over eighty percent of the universities in his survey pool introduced cross-disciplinary graduate programs.

The CTSA program has also put an emphasis on promoting multidisciplinary and provided support in making multidisciplinary teams with the hope that it can help research teams to come up with solutions for smooth flow from basic science and clinical science (Luke et al., 2015). NIH thinks that combining skills and approaches used in various disciplines can accelerate the translational process and help solve the puzzle of complex problems related to disease (Zerhouni, 2006). However, the relationship between receiving CTSA and increase of

[36] For instance, many research centers funded by NIH, which has much larger funding on the support for interdisciplinary research than NSF, such as Center for Evidence-Based Practice in the Under Deserved, Interdepartmental Neuroscience Center are labeled as doing interdisciplinary research (Aboelega et al., 2009). These centers try to integrate knowledge from various disciplines to broaden the scope of investigation and remove the potential roadblocks for the interdisciplinary collaboration (Aboelega et al., 2007). A more representative example of NSF's interdisciplinary center would be Engineering Research Center (ERC). ERCs are interdisciplinary, multi-institutional centers that links academic institutions, private companies and government to narrow the gap between academia and industrial engineering applications (National Science Foundation, n.d.-a). ERC program is similar to CTSA program in that it has the purpose of expanding the use of research outcomes from the academia. However, these two programs are different in several aspects. First, ERC focuses various fields such as security, health, energy and manufacturing (National Science Foundation, n.d.-a) whereas the CTSA program focusses on only field of life science. Second, while ERC program is associated with workforce in engineering, the CTSA program is more related to researchers in basic science and clinical science. Thirdly, ERCs are multi-institutional by nature (National Science Foundation, n.d.-a) but having partnering institutions is not mandatory for CTSA hubs.

multidisciplinarity of research is mixed. In the study by Luke et al. (2015), it was shown that the academic researchers who were part of a CTSA institution from the first place got more involved in cross-discipline collaboration than they used to after they joined CTSA related center. However, in their same study, Luke et al. (2015) showed that if they add the ones who joined the CTSA institution in the study sample, the overall tendency of cross-disciplinary collaboration decreases. The authors mention that there are two reasons for this trend. First, they claim that all publications, which were used to calculate the co-authorship trend, might not have been captured because of the time lag between funding and publication. Secondly, they maintain that the ones who join the institutions later would be junior researchers and they may tend to stick to their own discipline in their research. This is because what researchers need early in their career is recognition in their own field and they would try to make more publications on their own field.

The point I would like to make here is that the ultimate goal of CTSA is supporting research that produces translational outcomes, not multidisciplinary outcomes. Increasing multidisciplinary may help the researchers produce more translational outcomes but we have to keep in mind that those two goals are not the same. For the case of CTSA, the ultimate goal of the award is to produce non-clinical research outcomes that have clinical use or that links the research outcomes from basic science to the clinical side of the research pipeline. Increasing multidisciplinarity may have a positive impact in accomplishing this goal, but it is not same as achieving this goal. Therefore, multidisciplinarity should be considered as a mediator of the relationship between CTSA and translational research outcomes. In that regards, I propose the following hypothesis, which can be summarized as "Multidisciplinarity is a mean to reach the ultimate goal of translatedness." Figure 26 shows the depiction of this hypothesis.

Hypothesis 3. **The multidisciplinarity of institutions mediates the relationship between receiving CTSA and production of translational outcomes.**

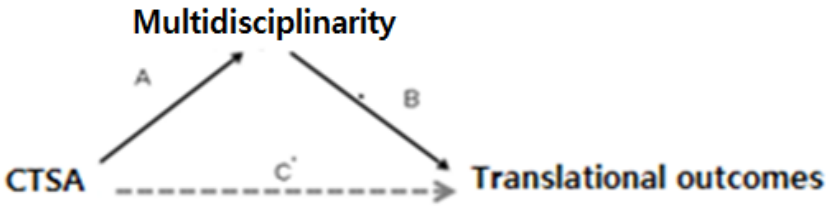


Figure 26. The depiction of Hypothesis 3

5.3. Data and methodology

5.3.1. *Empirical setting*

To test the hypotheses, I took institutions as the unit of analysis. As institution is the unit of analysis, the variables will be aggregated in institutional level. Some variables will use the sum of counts (e.g., number of publications with translational feature), whereas for some variables the average value will apply (e.g., multidisciplinarity).

I assume that the impact of the CTSA award appears after a few years and apply a three-year lag based on the finding by Ihli (2016). That is, in case an institution received the CTSA in year t, the impact of the CTSA award will be reflected in the articles that are published in year t+3. In this respect, the variables related to publications (e.g., collaboration density based on co-authorship, multidisciplinarity based on backward citation of publications) should be calculated in year t+3 to see the impact of the CTSA award in year t.

One disadvantage of using this approach is that the sample size may not be large, which equals the multiplication between the number of Carnegie R1 universities (115 institutions) and the number of years in the period of analysis (15 years), which is 1,725. However, the sample size is still much larger than what Cohen (1988) suggested as the necessary sample size to run a multiple regression, which is 764. The status of the institution receiving CTSA will be independent variable and the inter-organizational collaboration intensity and multidisciplinary score of the institutions will be mediating variables. The empirical setting of my analysis is depicted in Figure 27.

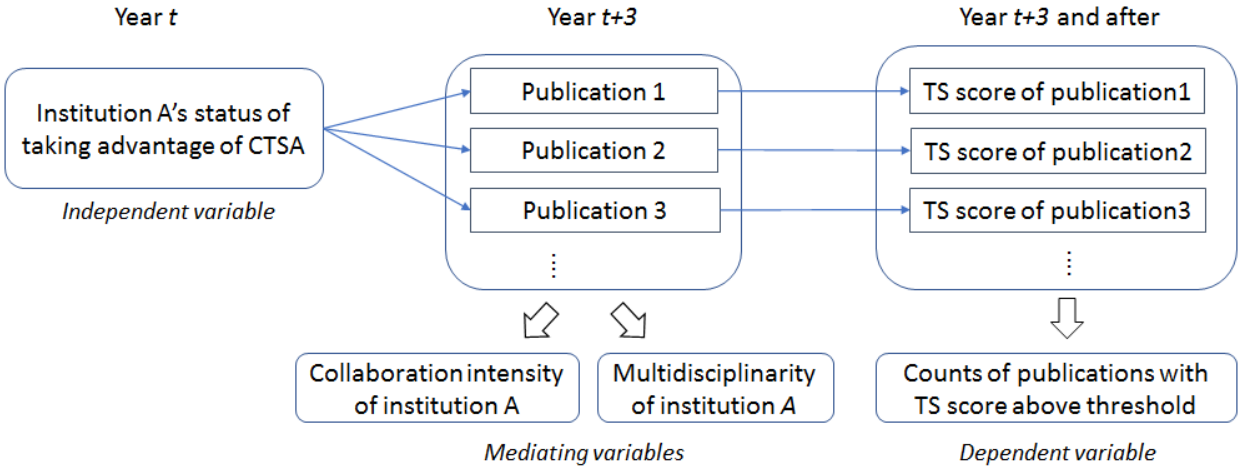


Figure 27. The empirical setting of my analysis in chapter 5

5.3.2. *Data and variables*

5.3.2.1. Dependent variable

The dependent variable of my analysis is the counts of journal articles that possess translational feature. After an article is published, we would be able to track its use in the

research community by looking at its forward citations (Llewellyn et al., 2019). The forward citation can take place from the year that an article is published and the total number of citation an article receives will increase over time. However, in contrast to the total counts of forward citation, the TS score will have a stable value after a certain period of time[37]. Considering this feature, I plan to use the TS score to check if an article has translational feature.

I plan to categorize an article as having a translational feature if TS score is above 0.150. This choice of the threshold is based on the validity tests presented in chapter 4. The validity tests showed that publications that are cited by patents had the average TS score value of 0.156 when prospective approach was applied, and 0.140 if retrospective approach was used. Also, among the four examples of basic science publications that result in drugs, the article by Rock et al. (1994) had the lowest TS score and the value was 0.149. Considering these values are all around 0.150, it seems reasonable to set the threshold value as 0.150. As the unit of analysis for this section is institution, the count of basic publications with TS score above 0.150 for each year will be used as the dependent variable. As the publication count variables are usually not normally distributed as in Gaughan et al. (2018), I will be using not only Ordinary Lease Square regression but also Poisson regression in my analysis.

5.3.2.2. Independent variable

The independent variable is a dummy variable on whether an institution received the CTSA award or not. I take the treatment group as the institutions that received the CTSA award and the comparison group as the institutions that didn't take advantage of the CTSA award. As in

[37] As presented in chapter 4, my reliability test shows that it takes around four years for a TS score to reach its stable value.

chapter 3, institutions that partnered with institutions receiving the CTSA award are considered as a member of the treated group. An institution that was selected as the CTSA receiving institution or partnered with the CTSA receiving institution at year *t* would have the value of 1 from year t. The institutions that never took advantage of the CTSA program will have zero value for the independent dummy throughout the period of analysis.

5.3.2.3. Mediating variables

Inter-organizational collaboration intensity[38]

The first mediating variable I will include is the intensity of collaboration. As discussed, many studies found out that collaboration has a significant impact on scholarly productivity (e.g., Lee & Bozeman, 2005). Centrality measures provide key characteristics of institutions regarding their position in the network (Hanneman & Riddle, 2005). Hence, I use centrality of the institution as the variable measuring inter-organizational collaboration. Among four centralities calculated in chapter 3, I will be using degree centrality as the first mediator of the model. Degree centrality is chosen as the CTSA program promotes direct relationship between institutions and it is what measures this feature.

Multidisciplinarity

One very big challenge in fostering multidisciplinary research is difficulty in defining and measuring multidisciplinarity precisely (Qin et al., 1997). Qin et al. (1997) gave a list of characteristics that multidisciplinary research possesses such as it having members using different approaches to solve problems, having members influenced by other members on

[38] More details on the measures related to inter-organizational collaboration can be found in chapter 3.

performing a task and integrating the analytical strength of disparate scientific disciplines. However, it is very hard to come up with an indicator that incorporated all these factors and there is no agreed-upon measure of multidisciplinarity (Qin et al., 1997). This leads to different measurements, assessment tools and evaluation processes (Wagner et al., 2011). However, from the list in Qin et al. (1997), we can see that knowledge integration is the must-include factor when defining or assessing multidisciplinarity of research (Wagner et al., 2011).

Though there is no unified measure on multidisciplinarity, scholars have been conducting research on multidisciplinarity using the measures they deem most appropriate. A large portion of studies use quantitative measure in measuring multidisciplinarity and bibliometric measures are the most developed and widely used quantitative measures (Ihli, 2016; Wagner et al., 2011). The simplest measure is assessing multidisciplinary research is the number of disciplines involved, which is based on the co-occurrence of author affiliations, keywords or headings (Schummer, 2004). One type of this co-occurrence approach is counting the papers co-authored by the people from more than one disciplines as in Shummer (2004). Schummer (2004) claimed that the co-occurrence approach could capture the cognitive and social aspects of relationship and exchange of knowledge between corresponding disciplines. However, this approach is not widely used as it is very tough to get affiliation information without errors (Wagner et al., 2011). As I mentioned in chapter 4, there is inconsistency in the information related to author's affiliation. There are differences in the information the author's affiliation contains, and the order in which the information is listed is very diverse. Also, we could not know exactly what discipline a researcher is engaged in just with the department name of his or her affiliation, as discussed in chapter 3.

Another widely used bibliometric measure is based on the citations of the paper (Wang et al., 2017). For instance, Wang et al. (2017) used the composition of citation categories to measure the multidisciplinarity of papers, which were later averaged by institutions in order to get the multidisciplinarity of a research system. Qin et al. (1997) used a similar method and counted the number of disciplines cited by a paper to calculate multidisciplinarity. To address the issue of difference in the size of the disciplines, relative measures or ranking measures are sometimes used (Wagner et al., 2011). In some cases, citations from disciplines outside the journal that a paper was published in are used as an indicator (e.g., Porter & Chubin, 1985). This approach is based on the assumption that when scientists cite sources outside their discipline, science-information interaction takes place (Qin et al., 1997).

All in all, we can see that there is no measure that can be used universally in all situations (Wagner et al., 2011). As in claim by Martin and Irvine (1983), among multiple sources and measures on multidisciplinarity, we need to make a selection on what is most appropriate based on the context of the research. Here, my choice on the measure of multidisciplinarity is average Shannon Diversity of publications based on the backward citation of each article averaged into institution-year level. The reason why Shannon diversity of backward citation of each publication is used is that it is essential to take into account citation categories when measuring multidisciplinarity of publications (Rafols et al., 2012; Wang et al., 2017). To calculate the Shannon Diversity of each publication, following equation will be used. Here, P_i is the portion of cited articles published in WOS$C_{i.}$

$$1 - \sum_i p_i log p_i, \tag{4}$$

After Shannon Diversity of each publication is calculated, it will be averaged into institutional-year unit. This approach is taken as individual publications can be considered as

elements of a research system and the average of a group of publication could provide unbiased measurement of multidisciplinarity (Wang et al., 2017).

5.3.2.4. Control variables

Research capacity (Researchers and R&D in life science)

One major factor that has a significant role in changing research performance of an institution is the research capacity of the institution. Research capacity of an organization, expressed as absorptive capacity in some contexts, affects the level of change in research performance due to change in collaboration level (Owen-Smith & Powell, 2003). The study by Owen-Smith and Powell (2003) included scientific capacity, which is measured as the publication counts, in their model measuring the relationship between collaboration and research patenting performance of academic institutions. We can expect that institutions having high research capacity would have higher research performance. Research capacity of an institution can be measured by research budget and the number of researchers (Cockburn & Henderson, 1998). In this regard, I will include both the size of biomedical research budget and number of researchers in life science as control variables.

Medical school

In biomedical research, whether an institution has a medical school or not also affects the relationship between collaboration and productivity (Owen-Smith & Powell, 2003). Owen-Smith and Powell (2003) claimed that research conducted in medical school is close to commercial needs and can contribute to increase in commercial products. In this regard, I included the existence of medical school as another control variable. If a university has a medical school, it is

more likely that the translational process will be smoother (Owen-Smith & Powell, 2003). As mentioned in study by Hicks and Katz (1996), hospital is an important entity in the process of biomedical research. There are also reports on the significant role of hospitals in biomedical research pipeline (Lander & Atkinson-Grosjean, 2011).

Institutional control

The last control variable I included is institutional control. Institutional control, whether an institution is public or private, has been studied widely (Dundar & Lewis, 1998). Some studies showed that being a private school lead to higher research productivity as private schools put more emphasis on research than teaching and other services (Dundar & Lewis, 1998; Jorand et al., 1988). As I'm taking one kind of research productivity as my dependent variable, though it is different with traditional measures like mere count of publication, the institutional control factor could have the potential to affect the result of my analysis. Table 14 shows the description of the variables that will be used in my analysis.

Table 14. Descriptions of the variables used in the analysis

Category	Key variables	Measurement strategy
Dependent variables	Translational publication count	Number of articles published in non-clinical science journal that has TS score higher than 0.150.
Independent variables	CTSA	Whether the institution took advantage of the CTSA award
Mediating variable	Intensity of collaboration	The frequency of collaboration with all other institutions measured by co-authorship (degree centrality)
	Multidisciplinarity	How multidisciplinary an intuition is based on the multidisciplinarity of articles it published
Control variables	Researcher	Sum of biomedical researchers in life science
	R&D	Size of the research budget in life science
	Medical school	Whether an institution has a medical school
	Private	Whether an institution is private school

5.3.3. *Method*

Difference-in-difference regression

Using longitudinal data with repeated measure can help us test the causal relationship between variables (VanderWeele & An, 2013), which will be used in my analysis. With the longitudinal data, I will be using multivariate regression that compares the values before the intervention and after the intervention. To do so, I will be using the difference-in-difference method. This approach may help address endogeneity to a certain degree. The empirical equation for the analysis is as follows:

$$Y_{i(t+s)} = \alpha + \beta \cdot CTSA_i + \gamma \cdot POST_{i(t)} + \delta \cdot CTSA_i \cdot POST_{i(t)} + X'_{it} \cdot \boldsymbol{\delta} + FE_i + FE_y + \varepsilon_{i(t)} \quad (5)$$

where *i* is a project and *t* is time.

$Y_{i(t+s)}$ is the count of translational publications for year *t*+*s*. Year *t*+*s*, not *t*, will be used as it takes some time from the start of the project to get the result of a project. This variable is a count variable that is not normally distributed as in Gaughan et al. (2018). Hence, I will test not only Ordinary Least Square (OLS) model, but also the Poisson regression model for the analysis. $CTSA_i$ is the dummy variable of my interest, which is the status of receiving CTSA.

$POST_{i(t)}$ is after-treatment dummy, which has a value of 1 if the period is after the treatment occurred. For the institutions in the control group which did not receive the CTSA award during the period of analysis, I need years that these institutions would have started receiving the CTSA award. I could just select a single year (e.g., 2008) and set this as hypothetical CTSA-starting year for all institutions that never took advantage of the CTSA

award. However, this approach would have minimal influence in reducing the selection bias problem. Hence, to get these hypothetical years, I used propensity score matching method. First, I looked at the of content of FOA of NIH's CTSA award (National Institutes of Health, 2017b). Then, I figured out factors that may influence NIH's selection of CTSA receiving institutions. With these factors, I calculate the propensity score of the institutions receiving CTSA using a logit model. Then, I matched CTSA non-receiving institutions and CTSA receiving institutions based on these scores using nearest neighbor approach. Then, I assigned hypothetical CTSA receiving year to the institutions that never took advantage of the CTSA award. The common support option was not applied as I need imaginary treatment year for all institutions. As a result, institutions that only have hypothetical CTSA receiving year will also have value of 1 for this variable when the period is after its hypothetical CTSA receiving year.

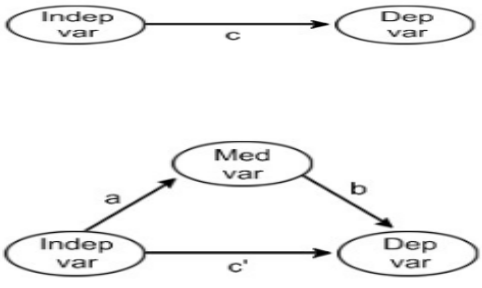
The interaction term $CTSA_i \cdot POST_{i(t)}$ is the variable of interest. If the coefficient of this term has significanlty positive value, we could maintain that the Average Treatment Effect (ATE) of the CTSA award on the increase of $Y_{i(t+s)}$ would be positive. As noted in previous section, various control variables (X'_{it}) (e.g., collaboration frequency, biomedical research budget) will also be used in the equation. Fixed effects on institutions (FE_i) and years (FE_y) will be applied in the equation to control for heterogeneity between institutions and years that is not captured by the control variables.

Mediation tests

To test the mediation effect, a method proposed by Baron and Kenny (1986) will be used with slight modification. In their seminal paper on mediation analysis, they introduced four conditions that have to be met for the mediation effect to be present. These conditions were based on the estimations from three regression models. First estimation is based on the result of

the model that regresses dependent variable on independent variable. From this model, we could get coefficient of independent variable, which is labeled as ***c***. The first condition for the mediation effect to be present is this coefficient ***c*** being significant. The second condition comes from the second model that regresses the mediator on the independent variable. The second required condition is that the coefficient of independent variable in this model, which is labeled as ***a***, should be significant. The third model regresses the dependent variable on both the independent variable and the mediator. This third model would give the coefficient of the independent variable, which is ***c'***, and the coefficient of the mediator, which is labeled as ***b***. The third requirement for the mediation effect to be present is that the absolute value of coefficient ***c'*** should be smaller than the absolute value of ***c***, which is from the first model that regress the dependent variable on the independent variable. The fourth condition is related to the indirect effect, which is $|c|-|c'|$ or $\mathbf{a \cdot b}$. Preacher and Hayes (2004) suggested that this value should be significant using the bootstrapping method. I used 10,000 iterations for the bootstrapping to see whether zero is not in the interval of the indirect effect value. All four conditions mentioned above, summarized in Table 15, should hold for the mediating role of the mediator to be present.

Table 15. Required conditions for medation effect to be present

Regression model	Condition
	• Condition 1: ***c*** should be significant • Condition 2 (optional): ***a*** should be significant • Condition 3: $\|c'\|$ should be smaller than $\|c\|$ • Condition 4: The value of indirect effect, which is $\|c\|-\|c'\|$ or $a \cdot b$, should be significant based on bootstrapping result (zero not in the interval)

5.4. Results

5.4.1. *Institution matching*

To calculate the propensity score on how likely an institution would take advantage of the CTSA award, I looked at the FOA of NIH's CTSA award (National Institutes of Health, 2017b) and identified four major criteria that may have impacted the selection of the CTSA recipients. Then, I linked these criteria with the factors that may have influenced the selection of the CTSA hubs.

The first criterion is related to the sharing of resource and research finding. In its FOA, NIH states that CTSA receiving institutions need to disseminate the solutions they found. In addition, collaborative leadership and communication are what NIH requires to the CTSA hubs. This means that the CTSA recipients should be in the positions that make them easy to act as a leader and share their resources with others. Hence, network centrality measure would be related in the assessment of the application and it should be included as one of the factors to calculate the propensity score of institutions taking advantage of the CTSA award.

There is one more reason that the network centrality measure should be included. The treated group in my analysis included institutions that officially partnered with CTSA hubs. We can easily expect that the CTSA hubs would establish an official partnership with the institutions that they have close relationship with. It is likely that the institutions with high value of network centrality would be the ones that have close relationship with many institutions, including CTSA hubs. Hence, having a high network centrality value would increase the probability of becoming a CTSA partnering institution. As the index of network centrality, I used degree centrality, which was calculated in chapter 3.

The second factor that may affect the selection of the CTSA institution is whether an institution has a medical school or not. NIH's FOA on the CTSA award requires (National Institutes of Health, 2017b) CTSA hubs to participate in research related to pediatric, geriatric and fields related to special population. As studies on these fields are conducted widely in hospitals and medical schools, it is likely that having a medical school would have positive impact of an institution being selected as a CTSA recipient. Therefore, I included the existence of a medical as one of the factors for calculating the propensity score and the information was obtained from NSF Higher Education Research and Development Survey (HERD).

The third potential factor is the racial diversity of the researchers. In the FOA, NIH encourages the institutions to diversify their students and faculty populations. Hence, having higher race diversity of people conducting research could have an influence on the selection of the CTSA award. To get the value for the race diversity, I used race composition of graduate students, post-doctoral researcher and non-faculty researchers in NSF Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS). I calculated Shannon diversity of the race composition, which was consisted of Hispanic, American Indian/Alaskan, Asian, Black, Native Hawaiian/Pacific and White.

The statement that is related to the fourth potential factor in the NIH's FOA (National Institutes of Health, 2017b) is "assembly of multidisciplinary translational teams." As mentioned in previous sections, multidisciplinary research has been promoted by government agencies like NIH and the CTSA award is not an exception. To obtain multidisciplinarity of institutions, the average multidisciplinarity of articles published by the institution was used. As explained previously, multidisciplinarity of each publication was calculated using Shannon diversity of

disciplines that appear in backward citation of publications. Table 16 shows the summary of factors that are potentially used by NIH when selecting CTSA receiving institutions.

Table 16. CTSA selection rules and potential factor affecting the selection

No.	Selection rules in NIH's FOA	Potential factors affecting the selection
1	"CTSA hubs are expected to develop, demonstrate, and then disseminate solutions to roadblocks to the efficiency and effectiveness of clinical and translational research. Plans for collaborative leadership and communication should consider the inclusion of a diverse range of internal and external partners and stakeholders."	Network centrality measure
2	"CTSA hubs should have plans to enhance the participation in research of pediatric, geriatric, and other special populations, and in some instances make them the focus of study. CTSA hubs should aim to include underserved populations, address health disparities, and approach cultural factors as a variable to be examined when developing translational innovations."	Having medical school
3	"This program encourages institutions to diversify their student and faculty populations and thus to enhance the participation of individuals currently underrepresented in the biomedical, clinical, behavioral and social sciences research enterprise, as described in NOT-OD-15-053."	Race diversity of the researchers
4	"Strategic goals under such a vision may include increased incentives for teamwork, facilitation of the assembly of multidisciplinary translational teams, promotion of collaborative efforts, and increased knowledge and awareness of what works best in team science."	Multidisciplinarity of organization's work

Source: FOA of NIH's CTSA award (National Institutes of Health, 2017b)

With four potential factors affecting the selection of the CTSA hubs, I calculated the propensity sore of all 115 Carnegie R1 universities using logistic regression. Then, I matched the institutions based on the nearest-neighbor option[39]. For some institutions, it was matched with more than one institution. For instance, Boston College, University of Texas at Arlington and Clemson University, which all had propensity score below 0.02 was matched with University of Wisconsin in Milwaukee, which has the lowest propensity score among the treated group (0.018). On the other hand, some institutions in the treated group are not matched with any of the institutions in the control group. This is because the highest propensity score of any institution in the control group is 0.792 and the institution with this score (University of Connecticut) is matched with its nearest neighbor, which is University of Utah with the propensity score value of 0.796. Hence, other institutions in the treatment group with propensity score larger than 0.796 are not matched with any other institutions in the control group. Table 17 shows the size of the groups based on the CTSA starting year, whether it is actual or assigned through institution matching.

Table 17. Group size based on CTSA starting year

Starting year	Actual CTSA	Hypothetical CTSA	Total
2006	9	5	14
2007	12	6	18
2008	14	20	34
2009	5	1	6
2010	9	22	31
2011	4	1	5
2012	1	1	2
2013	1	3	4

[39] Full list of institutions with matching results can be found in the Appendix F.

5.4.2. *Descriptive analysis*

Summary of the data

To get the sense of the characteristics of institutions, the summary of variables of two groups, the institutions that received the CTSA award (CTSA-ever) and the institutions that never received the CTSA award (CTSA-never), are presented in Table 18. For the data summary and all other analysis, only publications with five or more forward citations were considered when calculating the TS score to maximize reliability of the measure and results (See chapter 4.4.1.B for more details). In general, the CTSA-ever group has larger values for most of the variables than the CTSA-never group. For instance, the average number of non-clinical science publications per institution-year is 345.46 for the institutions in the CTSA-ever group while the value is 96.58 for the CTSA-never group. Differences also exist in the control variables like biomedical R&D expenditure and biomedical researcher counts. Hence, it is essential to include these variables as the covariates to help address omitted variable bias. One fortunate fact is that the variables related to forward citation of publications (e.g., mean forward citations, mean forward citations from clinical science and mean TS score) are not that different between two groups. For example, the number of forward citations a paper receives has an average value of 53.53 for the publications from the CTSA-ever group and 46.31 for the publications from the CTSA-never group. Also, the difference of the mean counts of forward citation from the clinical side is 2.88, which is not that large compared to the difference of the variables in the covariate group. As there is no big difference in this value between two groups, the average TS score values of two groups would not differ much.

Table 18. Data summary with comparison of CTSA-evers and CTSA-nevers[40]

Variable	Institutions that ultimately took advantage of CTSA					Institutions that never took advantage of CTSA				
	Obs	Mean	Std.dev	Min	Max	Obs	Mean	Std.dev	Min	Max
Non-clinical science publications	825	345.46	251.65	2	1659	900	96.58	76.25	1	424
Mean forward citations (per paper)	825	53.53	21.04	13.03	145.06	900	46.31	24.39	8.88	220.41
Mean forward citations from clinical science (per paper)	825	7.87	3.83	0.67	30.15	900	4.86	3.33	0	28.13
Mean TS score (institution-year unit)	825	0.13	0.04	0.03	0.23	900	0.10	0.05	0	0.27
R&D (million $)	825	310.29	202.11	2.51	882.69	900	93.47	78.50	0.90	373.77
Researcher (	825	17.09	11.15	0.60	85.11	900	8.23	5.05	0.75	26.61
Medical	825	0.83	0.38	0	1	900	0.34	0.47	0	1
Private	825	0.45	0.50	0	1	900	0.15	0.36	0	1
Degree centrality	825	171.46	41.08	14	228	900	113.62	50.63	4	220
Multidisciplinarity (institution-year unit)	825	2.02	0.12	1.60	2.35	900	25108	0.16	1.29	2.49
Race diversity	825	1.60	0.25	0.80	2.27	900	1.60	0.28	0.50	2.38

The trend of translational publication counts

For my analysis, the publication counts with TS score over certain threshold is used as the dependent variable. As I mentioned in previous section, a TS score of 0.150 will be used as the threshold for determining whether a publication has translational feature or not. To have better understanding of the value, I looked at the change of counts of non-clinical publications that have TS score above this threshold, which I call translational publications hereafter (Figure 28). In the figure, the average translational publication counts of institutions for the treated group (CTSA-ever) is indicated as solid line while the control group (CTSA-never) is indicated as dashed line. The shaded areas show the range of 25 and 75 percentiles values for each group. The year when the CTSA program started is 2006 is indicated by a vertical line on the graph. As we can see, the value for the CTSA-ever group is larger than the value of the CTSA-never group

[40] Data summary and correlations between variables having all Carnegie R1 universities in the sample are shown in Appendix G.

for all years. Also, CTSA-ever group is in an increasing trend while CTSA-never group is not. This finding suggests that not all institutions are producing more translational publication than before, but only the institutions that ultimately received the CTSA award are.

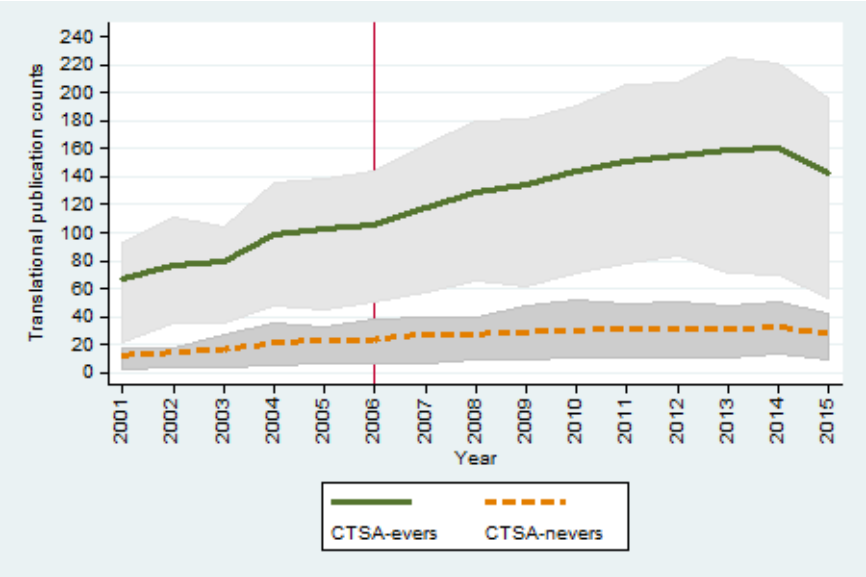


Figure 28. Trend of translational publication counts by calendar year

Figure 28 provides general picture on how the translational publication counts of institutions change over time, but it provides limited information on the impact of the CTSA award. This is because the year when each institution starts receiving the CTSA award differ. Hence, to get a better picture on how the CTSA award affects the change of the translational publication outcomes, I made a new trend graph taking years normalized by the years of CTSA receipt as the x-axis (Figure 29). Actual years of CTSA receipts were used as the reference year for the CTSA-ever institutions and hypothetical CTSA receipt year drawn from propensity score matching of institutions (see Appendix F for details) were used for institutions in CTSA-never group. From the graph, we can see that the treated group is in an increasing trend while the value

for the controlled group shows no increasing trend. Table 19 shows the mean translational publication counts before and after the receipt of the CTSA for both the treated and the controlled group. The number of translational publications that institutions in the treated group produced increased by 59.84, whereas the controlled group only increased by 8.42. This means that the change of mean translational publication counts at the institutions that received the CTSA award is 51.42 larger than the change of the mean production of translational publications by the institutions that never received the CTSA award.

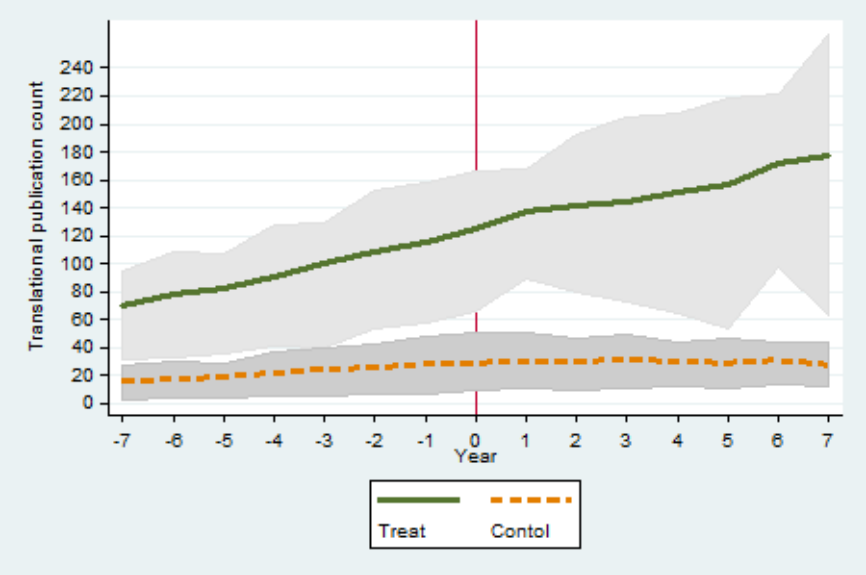


Figure 29. Trend of translational publication counts based on institution-matching without caliper option

Table 19. Mean translational publication counts before and after CTSA for all institutions

	Pre-CTSA	**Post-CTSA**	**Difference**
CTSA-ever group	94.38	154.22	59.84
CTSA-never group	21.98	30.40	8.42
		Difference-in-difference value	51.42

One thing to note in Figure 29 is that the increasing trend of translational publication counts is present even before the CTSA award starts to appear. This may suggest that there may be systematic difference between the treated and the comparison group. This result was predictable beforehand as the top-tier universities in biomedical science are also CTSA recipients and there may not be good comparable institutions in the CTSA-never group. To address this concern, I went further and applied a more stringent matching method, which is applying caliper value of 0.05 in the matching process. This means that if the value of propensity score of an institution differs from propensity score of any other institutions in the comparison group by 0.05 or higher, that institution is not matched with any other institution and not used in the analysis. Dropping unmatched institutions would provide us with comparison group more similar to the treatment group and give me more ability to make the causal claim about the impact of the CTSA award.

If I apply the caliper option, 24 institutions are dropped, of which most of them are the institutions with high propensity scores[41]. These institutions, who are in the CTSA-ever group are dropped as they don't have any comparable institutions that have high enough propensity scores as them. Dropping these institutions in the analysis would, on the one hand, reduced the size of the sample. However, the total number of observations, which is 1080, is still much larger than 764 observations that are needed to run multivariate regression (Cohen, 1988). Figure 30 shows the trend of translational publication counts by year after applying new institutional matching result. We can see that the pre-treatment trend line of two groups got more parallel. This could mean that the institution-matching method was somewhat more effective and the new treatment group and the new comparison group are now more similar to each other. In this

[41] Few examples are Emory University, Johns Hopkins University and Harvard University, which are all leaders in biomedical research in terms of research capacity and research performance.

respect, I will use the model with this matching method as one of the models in my analysis. One more thing to note is that there is a sudden rise of the translational publication counts for the treatment group right after the CTSA award come into effect, whereas there is no notable change happening for the control group. Table 20 shows the mean translational publication counts before and after CTSA recipient based on the new institutional matching method. When compared with the values in Table 19, the change of the pre-post difference reduced substantially for the treated group (24.47), while there is only slight change for the control group (0.12). This is because 23 universities with very high expected propensity of receiving the CTSA have been dropped for the new analysis and these institutions produce large number of translational publications. Even without these top-tier universities, the difference-in-difference value of translational publication counts is positive with the value of 26.81.

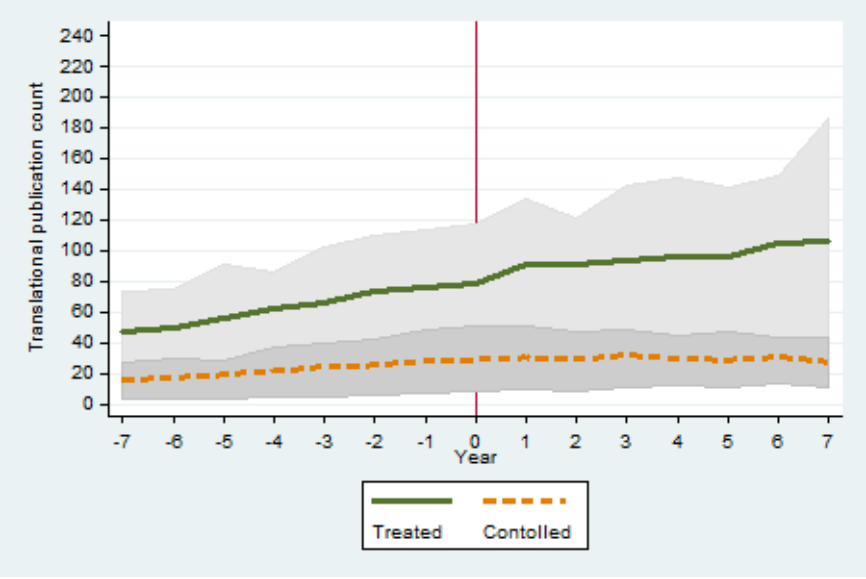


Figure 30. Trend of translational publication counts based on institution-matching with caliper 0.05 option

Table 20. Mean translational publication counts before and after CTSA based on institution matching with 0.5 caliper option

	Pre-CTSA	**Post-CTSA**	**Difference**
CTSA-ever group	62.34	97.45	35.11
CTSA-never group	22.05	30.35	8.30
		Difference-in-difference value	26.81

5.4.3. *Estimation results*

Test on Hypothesis 1: DID regression

Table 21 shows the estimations of the difference-difference regression models based on propensity score matching without caliper option[42]. The result table shows two extreme cases, the ones without any control variables and the ones with all control variables. Both the result of the model based on OLS regression and Poisson regression is shown. Regarding fixed effects, models either have only year fixed effect or both year fixed effect and institutional fixed effect. Year fixed effect is included in all models as the growth of the support for translational research differs by year, mostly in an increasing trend, and other variables have limited power capturing this change.

The most important thing to note is that the difference-in-difference variables (**CTSA X POST**) are all positive at 0.01 significance level. The value of the coefficients gets smaller as control variables are included for the OLS models. It means that the control variables also have explanatory powers in explaining the difference of translational publication counts between the treated group and the controlled group. Another thing to note is that the coefficients of **CTSA** are

[42] As applying caliper in institutional matching only gives us the ability to explain smaller set of comparable CTSA recipients, estimation of models without caliper option in institutional matching were also conducted. The results are shown in Appendix H and these model estimations generally have similar results with the result shown in main text, which is a sign of robustness of the result in the main text.

also positive with significance in models without institutional fixed effects. This means that being a CTSA receiving institution itself leads to a greater number of publications with translational feature. Using model 7, for example, we can see that CTSA recipients have approximately 2.20 (= exp(0.808)) more translational publications per year before the impact of the CTSA award come into effect and 2.32 (=exp(0.045+0.808)) more translational publications during the post-treatment period, holding all other variables constant. One more important thing to note is that the coefficients of **POST** are not significant when all control variables are included for the OLS model. This means that the translational publication counts for institutions not receiving the CTSA award does not experience an increase of translational publications after the hypothetical CTSA receiving year, which is confirmed by the trend graphs shown above (Figure 28, 29 and 30).

If we look at coefficients of the control variables, we can see that the researcher counts (**Researcher**) and R&D expenditure (**R&D**) have significant impact on the translational publication counts, although the impact of R&D expenditure is minimal compared to other factors. One interesting thing to note is that having medical school (**Medical**) has mixed effects. This may be interpreted as the universities with medical schools doesn't always lead to more production of non-clinical publications with translational feature. This finding could indicate that schools with medical school produce clinical papers, rather than creating basic science publications, in the first place, which are not captured in the dependent variable. Another interesting fact is that being a private school (**Private**) also shows mixed results. It has significantly positive impact in OLS model, but insignificant effect for the Poisson model. Therefore, we couldn't make clear claim whether being private school leads to more production of translational research outcome.

Table 21. Estimation of main difference-in-difference regression models (dropping unmatched institutions)

	OLS		Poisson		OLS		Poisson	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CTSA X POST	13.71***	13.65***	0.04*	0.04*	8.63***	8.70***	0.05*	0.05*
CTSA	50.66***	-	1.07***	-	25.64***	-	0.81***	-
POST	-3.72*	-3.77**	-0.03	-0.03	-1.58	-1.69	-0.04	-0.04
Researcher					7.29***	5.50**	0.05*	0.04
R&D					0.17***	0.15***	0.00***	0.00*
Med school					1.39	-6.67*	0.19**	0.07
Private					17.73***	-	0.01	-
Multidisciplinarity					15.69**	15.09**	1.00***	1.01***
Centrality					0.08***	0.04	0.00***	0.00***
Constant	19.14***	36.59***	3.09***		-34.54***	-9.09	0.58*	
Observation	1080	1080	1080	1080	1080	1080	1080	1080
R^2	0.342	0.115			0.736	0.615		
Pseudo R^2			0.019	0.137			0.074	0.153
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Note: *** p < 0.01, ** p < 0.05, * p < 0.10

With consistency in the significance and sign of the DID variable's coefficient, we may assert that the CTSA award did lead to the increase production of basic science publications with translational feature. There may be various reasons for this increase. One thing we can think of is the increase of publications that got direct support from the CTSA award. These are the publications that acknowledge the CTSA award's grant number on their papers, which probably have high TS scores. Figure 31 shows the trend of average TS score of three groups of papers, which are publications from CTSA-never group, publications from CTSA-ever group and publications that acknowledge direct support from the CTSA award. On average, publications that got support from the CTSA award have larger TS score than other two groups of publications. If we check the portion of publications that acknowledge the CTSA award, we can easily check that it is in increasing trend (Figure 32). In some universities such as University of

Chicago and University of Cincinnati, the portion increase to the value higher than 20% in 2015 (see Appendix I for more details).

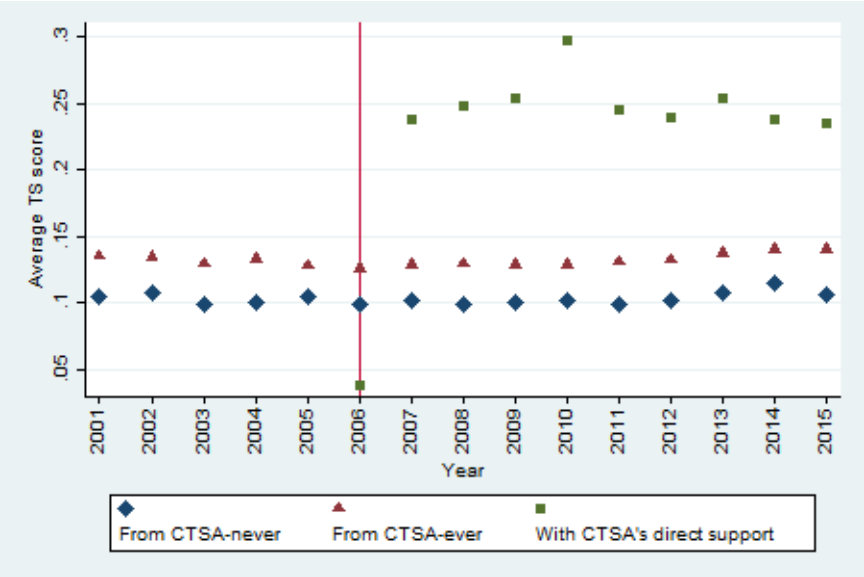


Figure 31. Average TS score by publication group

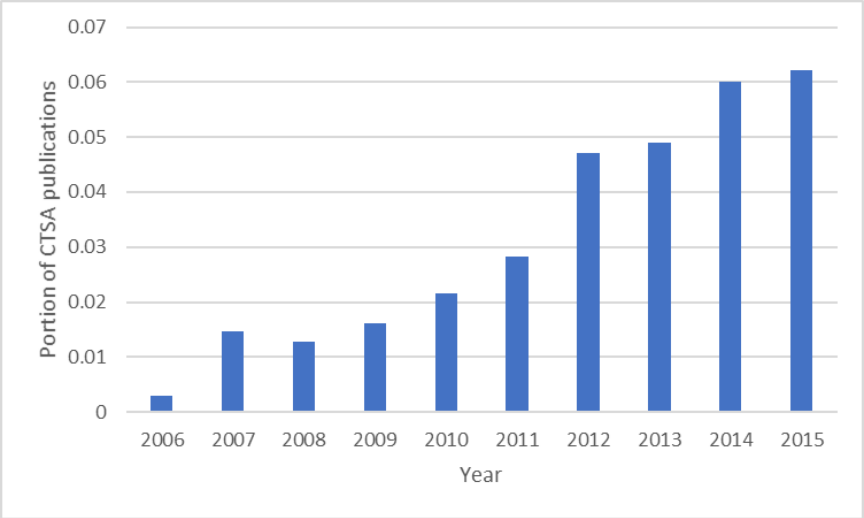


Figure 32. Change of the portion of CTSA award acknowledging publications by year

To examine this question, I identified all publications that received direct support from the CTSA award and, thus, might, have higher TS scores. This job was done to conduct new analyses that do not include these publications in the dependent variable, which could help us have better understanding of the impact of the CTSA award. If the estimations from the new models show similar results with the previous ones, we may claim that the CTSA award had an impact across the whole biomedical research environment of the universities and even changed the nature of publications that are not directly affected by the CTSA award. If the results find that there is no significant difference-in-difference effect when excluding CTSA supported articles, we may conclude that the CTSA award only changed the aspects of the publications that got direct support from the award but was not making the institution broadly more translational.

Table 22 shows estimation results of difference-in-difference Poisson regression models excluding publications with direct CTSA support. We can see that the coefficients of the difference-in-difference variables (**CTSA X POST**) now become insignificant. This finding suggests that the impact of the CTSA award is only valid on the publications that received direct support from the award. Therefore, we would not be able to make a claim that the CTSA award changed the whole institutions in a way that also influenced studies that didn't get its support.

Table 22. Estimation of difference-in-difference regression models excluding publications with direct CTSA support

	(1)	(2)	(3)	(4)
CTSA X POST	-0.03	-0.03	0.00	0.00
CTSA	1.06***		0.82***	
POST	-0.01	-0.01	-0.02	-0.02
Researcher			0.06**	0.05*
R&D			0.00	0.00
Med school			0.19**	0.08
Private			0.00	
Multidisciplinarity			1.06***	1.07***
Centrality			0.00***	0.00***
Constant	3.10***		0.48	
Observation	1080	1080	1080	1080
Pseudo R2	0.016	0.109	0.079	0.127
Year FE	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES

Note: *** p < 0.01, ** p < 0.05, * p < 0.10

Test on Hypothesis 2: Mediation test on inter-organizational collaboration

Table 23 shows the summary of the estimation results that take degree centrality of institutions based on inter-organizational collaboration as the mediator. The tests were conducted separately by year with 3-year lagged values applied for the treatment dummy (**CTSA**) and other control variables (**Researchers, R&D, Medical, Private**). To confirm that the potential mediator takes the mediating role, all four conditions must be satisfied. We can see that the mediation effect is not present in the early years from 2009 to 2013. This is mainly because condition 2 is not satisfied in these years, which means that the impact of the CTSA award on the degree centrality is not significant during this period. This leads to the situation where indirect effect not having a significant value and condition 4 not being satisfied.

On the other hand, when it comes to year 2014 and 2015, the mediating role of the degree centrality of institutions starts to appear. This finding may indicate that the inter-organizational collaboration rate starts mediating the relationship between the CTSA award and translational publication counts when the CTSA program finds its feet. We could expect that it would take some time to stack the inter-organization collaborations that would really help the increase of translational outcomes and this is confirmed by my findings.

Table 23. Result of mediation tests taking inter-organizational collaboration as mediator

Year	Individual conditions				Existence of mediation effect	Portion of indirect effect to the total effect
	Condition 1	Condition 2	Condition 3	Condition 4		
2009	YES	NO	YES	NO	No	NA
2010	YES	YES	YES	NO	No	NA
2011	YES	NO	YES	NO	No	NA
2012	YES	NO	NO	NO	No	NA
2013	YES	YES	YES	NO	No	NA
2014	YES	YES	YES	YES	Yes	.0017
2015	YES	YES	YES	YES	Yes	.0022

Note: More detailed result with actual value of coefficients can be found in Appendix J

Test on Hypothesis 3: Mediation test on multidisciplinarity of institutions' publications

Table 24 shows the result of mediation tests taking multidisciplinarity as meditators from 2009 to 2015. As shown, condition 2 is not satisfied for all years, which means the CTSA award doesn't impact the change of multidisciplinarity score of institutions significantly. This may be because there is no clear difference of multidisciplinary of institutions between the institutions

that took advantage of the CTSA award and the ones that are not affected by the CTSA award (Figure 33).

Table 24. Result of mediation tests taking multidisciplinarity as mediator

Year	Individual conditions				Existence of mediation effect	Portion of indirect effect to the total effect
	Condition 1	Condition 2	Condition 3	Condition 4		
2009	YES	NO	YES	NO	No	NA
2010	YES	NO	YES	NO	No	NA
2011	YES	NO	NO	NO	No	NA
2012	YES	NO	NO	NO	No	NA
2013	YES	NO	NO	NO	No	NA
2014	YES	NO	NO	NO	No	NA
2015	YES	NO	NO	NO	No	NA

Note: More detailed result with actual number can be found in Appendix J

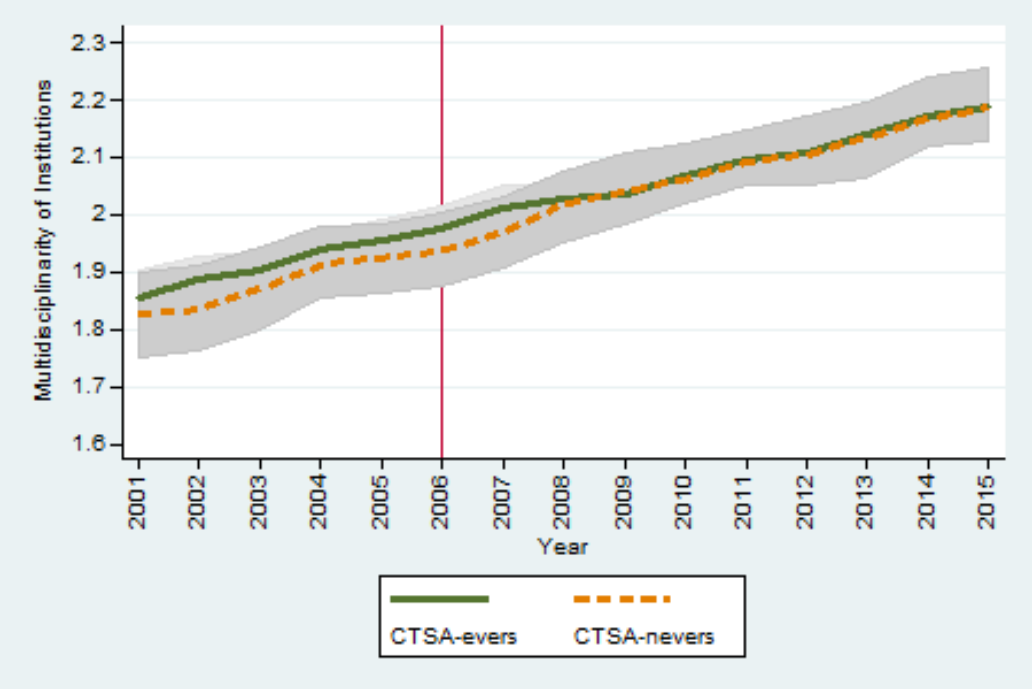


Figure 33. Trend of multidisciplinary score of institutions

5.5. Conclusion

In this chapter, I assessed the impact of the CTSA award by taking the number of translational publications as the outcome. A publication was considered as a translational publication if it has the TS score above 0.150, the threshold value that is drawn from my validity tests of the measure. Results from difference-in-difference regression models with various specifications showed that the CTSA award leads to increase of translational publication counts in institution level. However, it was shown that the CTSA program only affected the publications that got direct support from the program but didn't change entire institutions towards more translational research. This may seem like a limitation of the CTSA award, but if we see it in a different way, this could mean that the award has a substantial effect on the affected publications. Hence, if the program spreads throughout the institutions and more researchers start taking advantage of the program, it may change whole organization's characteristics related to translational research.

Mediation tests of two factors that NIH emphasizes for the successful operation of the CTSA program, which are inter-organizational collaboration rate and multidisciplinary of institutions, showed different results. It was shown that the inter-organizational collaboration rate, measured as degree centrality of the institutions, mediates the relationship between the CTSA award and translational publication counts five years after the CTSA program started. The main reason why inter-organizational collaboration rate did not have mediation effect during the early period was that the CTSA award did not lead to the change of the degree centrality of institutions. This result suggests that the CTSA award did not lead to the creation of collaboration that had

positive impact on the translational research at the beginning of the CTSA program. The newly created links might not have brought helpful resources to the institutions such as complementary skills or access to useful equipment (Bozeman and Corely, 2004; Katz & Martin, 1997; Gaughan et al., 2018; Owen-Smith & Powell, 2003).

The mediating role of the institution's multidisciplinarity did not appear during the whole period of analysis. Again, the reason why the mediation effect is not present is that the CTSA award did not lead to the change of multidisciplinarity of institutions. In my analysis, the multidisciplinarity of institutions is calculated based on the composition of backward citations of publications. Hence, CTSA award having no impact on the multidisciplinarity of institutions can be interpreted as CTSA award not having a significant impact on changing researcher's citing patterns. The trend of multidisciplinarity score of institutions is increasing (Figure 33), maybe due to the positive perceptions of the multidisciplinary research (Baumwol et al., 2011; Van Rijnsoever & Hessels, 2011) and the need of knowledge from various fields for answering complex problems (Aboelela et al., 2007). However, the CTSA award was not the main factor that caused the change. The online tools that the CTSA program provided with the aim to promote multidisciplinary research (e.g., Profiles Research Network Software) might not have worked well. Rather, it is possible that these tools might have only helped the researchers to find partners in their own disciplines, which is claimed by Van Alstyne and Brynjosfsson (1996).

There are some limitations of my analysis that is used in this chapter. First, as mentioned previously, the selection bias issue is not fully addressed. The institutions that were selected as the CTSA award recipients could be the ones that put more effort on translational works and might have increased their translational outcomes even without the CTSA award. I tried to alleviate this issue by using difference-difference regression after matching institutions,

assuming lagged effect of the CTSA award, dropping the institutions with very high propensity scores of receiving the CTSA award to check robustness of results and so forth. However, due to the nature of quasi-experimental research design setting as my study, the endogeneity problem may still be present. To address the endogeneity problem further, I may need to apply more precise institutional matching technique. Getting to know the real factors affecting NIH's evaluation on the selection of CTSA recipients or the actual evaluation scores of each institution during the NIH's selection process would help me get more accurate matching of institutions.

CHAPTER 6. DISCUSSION

6.1. Summary of findings

With the aim to bridge "the valley of death" in biomedicine, NIH has been increasing its support for translational research since the early 2000s. A clear indication of this effort is the creation of the CTSA award that provides infrastructure with multiple uses, tools for collaboration, education and administrative support and other resources to facilitate translational research. Since the CTSA program has been in operation for over a decade, it is important to understand whether and to what extent the program has achieved its intended goal. In this respect, I first examined whether the award had a positive impact on increasing collaboration among institutions, which is one of the major factors that NIH emphasizes for the successful operation of the program (National Institutes of Health, 2017b). In addition, I analyzed whether the CTSA awards helped institutions publish more articles that were translational in nature, another key goal of the program (Han et al., 2018).

My first analysis focused on the impact of the CTSA award on the collaboration network among Carnegie R1 universities, which are academic institutions with high research capacity. Using co-appearance on the author affiliation list of publications as evidence of inter-organization collaboration, I calculated various measures related to the collaboration network of biomedical research. These measures were used in my descriptive analyses and as variables in network regression model (e.g., ERGM) estimations. The results showed that the biomedical research collaboration is becoming denser and less centralized as the impact of the CTSA award gets realized. Institutions that received CTSA awards had more links with other Carnegie R1 universities than institutions that did not receive CTSA awards. However, the findings suggest

that the institutions that took advantage of the CTSA program are not interacting more with the institutions that never took advantage of the CTSA award. Rather, these institutions increase their collaboration with only the ones that ultimately took advantage of the CTSA award. This finding may not correspond precisely with NIH goals, as NIH would prefer major universities to have a wide range of partners, not just other elite institutions (National Institutes of Health, 2017b).

My second analysis examined the relationship between the CTSA award and the change of the translational publication counts. I aimed to examine the relationship between the CTSA award and the production of translational publications and see if two major factors that the CTSA award emphasizes (inter-organizational collaboration and multidisciplinary research) mediate the relationship. A unique measure – TS Score – was proposed and used to assess whether a publication was translational or not. The TS score is defined as the portion of forward citations from clinical science among all forward citations that a non-clinical publication receives. A new measure was proposed as there was no easy-to-use measure widely used that fits my research design. The TS score was shown to be reliable if a publication received five or more forward citations. Also, several tests using a group of publications were conducted to test the validity of the measure.

To classify a publication into publications with translational feature and one without, I had to set the threshold TS score value. Based on the results of validity tests, I selected 0.150 as the threshold value, which means that I considered publications with TS score higher than 0.150 as translational publications. As my unit of analysis was institution, my dependent variable was the number of non-clinical science publications with TS score larger than 0.150. The result of difference-in-difference regression with institution matching showed that there is a direct impact

of the CTSA award on the increase of translational research outcomes. However, this impact disappeared if the counts of publications that got direct support from the CTSA award were not included in the counts of the publications. This suggests that the CTSA award had changed the nature of the publications that are influenced by the award, but it didn't change the nature of biomedical research at these institutions more broadly.

Regarding the mediating effect, at the beginning of the CTSA award, neither the rate of inter-organizational collaboration rate nor the multidisciplinarity of institution's publications had significant mediating role on the relationship between the CTSA award and the translational publication counts. However, inter-organizational collaboration rate started mediating the relationship between these two factors five years after the CTSA program started. This finding may indicate that institutions are starting to create inter-organizational collaboration in a way that may help their translational research after the program is well established.

6.2. Implications

The first analysis of this dissertation shows how the collaboration network of biomedical research has been changing over time. I found that receiving the CTSA award is associated with increasing the number of collaborators and the rate of collaboration. This result may suggest that the CTSA award brought helpful resources to the institutions that took advantage of the award. According to the resource dependency theory, more resources could lead institutions to have more collaborations, which is confirmed by my analysis. Regarding the theory of self-interest paradigm, my result only partly supports it. The CTSA award did lead institutions to seek more partners but it turned out that they mainly focus on expanding collaboration with high-tier

universities. This could mean that benefit that can be gained by collaborating more with low-tier universities, which is future financial benefit provided by NIH by being selected as a CTSA recipient, is smaller than the potential benefits that can be obtained by working with higher-level universities. However, the result from my second analysis indicates that even this kind of collaboration pattern change didn't lead to an increase of translational publications until recently.

The findings from my second analysis show that the CTSA award led to an increase in the number of translational outcomes, which was the result that NIH desired. However, the result is not in the form that NIH really wanted. My results showed that the CTSA award only influenced the publications that got direct support from the award. Even if a group of publications were produced in the institutions that took advantage of the CTSA award, their nature of translatedness didn't change much if it is not directly affected by the award. This means that the CTSA program failed to change the research environment of entire institutions toward more translational research.

The limited impact of the CTSA award could be interpreted in two ways. First, the impact of the CTSA award may not be large enough to induce broader systematic change. Though the absolute size may be large, the portion of the grant among total research expenditures at recipient universities may not be large enough to influence the whole institution. For instance, of the total sponsored funding of Emory University in 2016, only about 6% were related to the CTSA program (Georgia Clinical & Translational Science Alliance, n.d.). Also, as shown in previous chapter, the average portion of publications that acknowledge the CTSA award on their publications is only around 6% in 2015 (Figure 32). The second explanation for the limited effect of the CTSA award could be related to lack of time for the impact of the CTSA award to appear. It is true that it has been more than ten years since the CTSA program started,

but it is possible that we need more time to see the impact of the CTSA award to appear. It may take a great deal of time for the factors related to the change of translational publication counts to change. As noted previously, it took five years to change the collaboration patterns of institutions toward a form that could help translational research. Other factors may also require as much or more time to change to a form that could help fasten the translational process.

To overcome this limitation, various measures can be made. First, CTSA recipients could try to spread the program more widely within the institution and introduce the resources from the CTSA program more to the researchers. If more researchers get to know about the CTSA award, the number of studies getting direct support from the CTSA award will increases. As shown, the publications that received direct support from the CTSA award have high TS scores on average. Therefore, the number and portion of translational publications, based on my definition using the TS score, will increase as more studies get direct support from the CTSA award. Also, more use of the CTSA program by scientists may change the whole institutional research culture towards more translational research friendly environment to a certain degree, which is the form that NIH ultimately wants (Han et al., 2018). The second approach could be increasing the size of the workforce who specialize in translational research. Education and training programs are major parts of the CTSA program (Weber, 2013) and the current approach of training researchers related to translational research is focused on educating acting clinical scientists (Weber, 2013; Luke et al., 2013). However, it is very hard to bring scientists in basic science silo and scientists in clinical science silo together (Butler, 2008; Blümel, 2017). Hence, even a scientist in particular discipline have some knowledge in another discipline, he or she may not collaborate with people in another discipline. If it is hard to make scientist from two groups work together, it may be effective to train researchers who can dedicate their work to tasks related to translation of

outcomes from the non-clinical science into clinical setting. If the number of dedicated researchers in translational research increases, the translational process can proceed faster even if scientists specialized in basic science and clinical science stay in their own field

6.3. Limitation and suggestion for further studies

Though the values of the coefficients in model estimations are statistically significant and the result supports the proposed hypotheses, my study, like all research, has room for improvement. In my concluding section, I would like to propose some approaches that I or others could use to confirm, strengthen confidence in, and build upon my results. First of all, I could not make a strong claim that endogeneity problem is fully addressed. It is possible that the organizations receiving CTSA may be the ones that will be more connected with other organizations regardless of the CTSA award. As these organizations are mostly high-tier organizations in the field of biomedicine, they could have received more collaboration request from their potential partners than the organizations that didn't receive the CTSA award. Also, these institutions could be the ones that are putting more effort on increasing translational works than the institutions not affected by the CTSA award. This may have caused selection bias of the samples and resulted in the bias on the estimators. I tried to alleviate this issue by matching institutions and applying difference-in-difference regression method. However, there is still more that can be done for the improvement.

One of the ways to alleviate the endogeneity issue further is developing a better comparison group. In case I know the factors that actually affect NIH's selection on CTSA recipients, I would get more precise propensity score of institutions and this could lead to better

institutional matching results. Also, I may also find a good instrumental variable on the list, which is the variable that affects the selection of the CTSA receiving institutions but doesn't affect the change of my dependent variables (e.g., collaboration rate, translational publication counts). In this case, I could adopt two-stage least squares (2SLS) regression analysis, which has a profound effect on solving the endogeneity problem. What is more helpful in getting better institutional matching result is getting a list of institutions that applied for the CTSA program each year. The institutions that applied but failed to get the award can be classified as comparative institutions to the CTSA recipients for the specific year of application. It would be even better if I get to know what scores each institution received during NIH's evaluation process regarding the selection of CTSA recipients. In this case, I could do institutional matching by pairing institutions that have close scores. If the number of institutions around the threshold score point that determines the selection of CTSA receiving institutions are large enough, I may also apply regression discontinuity method for the analysis.

Regarding the mediation analysis, one major limitation is that I conducted year-by-year analysis rather than applying longitudinal perspective. My approach of applying lags between the impact and the effect could avoid reverse causality between variables to a certain degree, but the estimators in this cross-sectional setting could still be biased (Maxwell, Cole, & Mitchell, 2011). To conduct longitudinal analysis on mediation, other methods such as Structural Equation Modelling (SEM) should be used (Maxwell et al., 2011). SEM can apply longitudinal framework and it can be used for causal inference in various settings (Maxwell et al., 2011). Thus, for further study, I may adopt longitudinal SEM to get more reliable results on the mediation of two major factors that NIH puts a lot of emphasis on.

Finally, conducting more robustness tests could help increase the reliability of the results from my analyses. I've tried to test my results with various settings including dropping unmatched institutions and using different dependent variables. However, performing more analyses under additional model specifications could increase the robustness of my results. First, I may test my result with other settings regarding time-related variables. I may apply different time interval between the impact and effect, which is three years in my current setting. Also, to test the placebo effect of the CTSA award, different year can be applied for the years of treatment. Secondly, I could use other values for the threshold TS score that is used in classifying publications into translational publications and non-translational publications. My current approach is using the threshold value of 0.150, which is based on my validity tests of the TS score in a small group of publications. However, it is possible that I get different threshold value if I conduct the validity tests with another sample of publications. In the process, I should keep in mind that testing with a larger sample of publications can increase the reliability of the threshold value.

APPENDIX A. List of CTSA hubs and their partnering institutions

Start year	CTSA-lead institution (CTSA-hub)	CTSA-partnering institution
2006	**Columbia University***	
	Duke University*	
	Mayo Clinic	
	Oregon Health & Science University	Kaiser Permanente's Center for Health Research
	Rockefeller University	
	University of California At Davis*	
	University of California San Francisco	
	University of Pennsylvania*	The Children's Hospital of Philadelphia, the Wistar Institute, the University of the Sciences in Philadelphia
	University of Pittsburgh*	**Carnegie Melon University***
	University of Rochester*	
	University of Texas Health Science Center at Houston	
	Yale University*	
2007	**Case Western Reserve University***	
	Emory University*	**Georgia Institute of Technology*,** University of Georgia* (partnership started in 2017), Morehouse School Medicine
	Johns Hopkins University*	
	University of Chicago*	Rush University, Advocate Health Care, Loyola University, Illinois Institute of Technology
	University of Iowa*	
	University of Michigan At Ann Arbor*	
	University of Texas Sw Med Ctr/Dallas	
	University of Washington*	Fred Hutchinson Cancer Research Center, Seattle Children's Hospital
	University of Wisconsin Madison*	Marshfield Research Clinic
	Vanderbilt University*	Meharry Medical College
	Washington University in St. Louis*	**University of Missouri***, St Louis University, BJC healthcare
	Weill Medical Coll of Cornell Univ*	Memorial Sloan-Kettering Cancer Center, the Hospital for Special Surgery, New York-Presbyterian Hospital, Hunter College
2008	Albert Einstein College of Medicine	
	Boston University*	
	Harvard University*	
	Indiana University*	**Purdue University*, the University of Notre Dame***
	Northwestern University*	
	Ohio State University*	
	Scripps Research Institute	
	Stanford University*	
	Tufts University Boston*	**Massachusetts Institute of Technology*,** Rand Corporation
	University of Alabama at Birmingham*	
	University of Colorado Denver	**Colorado State University***
	University of North Carolina Chapel Hill*	North Carolina State University* (partnership started in 2018), RTI International, North Carolina Agricultural and Technical State University

	University of Texas Hlth Sci Ctr San Ant	
	University of Utah*	
2009	Icahn School of Medicine at Mount Sinai	
	Medical University of South Carolina	
	New York University School of Medicine*	
	University of Arkansas Med Scis Ltl Rock*	
	University of Cincinnati*	
	University of Florida*	Florida State University* (partnership started in 2015)
	University of Illinois at Chicago*	
	University of Texas Medical Br Galveston	
2010	**George Washington University*** (Children's National Medical Center)	
	Georgetown University*	Howard University, MedStar Health Research Institute, Oak Ridge National Laboratory, and the Washington Veteran's Affairs Medical Center
	Medical College of Wisconsin	**University of Wisconsin-Milwaukee*,** Bloodcenter of Wisconsin, Children's Hospital of Wisconsin, Froedtert Hospital, Marquette University, Milwaukee School of Engineering, Milwaukee VA Medical Center
	University of California Irvine*	
	University of California San Diego*	
	Univ of Massachusetts*	
	University of New Mexico*	
	University of Southern California*	
	Virginia Commonwealth University*	
2011	**Pennsylvania State Univ***	
	University of California Los Angeles*	Cedars-Sinai Medical Center, Charles R. Drew University of Medicine and Science
	University of Kansas*	Kansas City University of Medicine and Biosciences, St. Luke's Health System/University of Missouri-Kansas City
	University of Kentucky*	
	University of Minnesota Twin Cities*	
2012	**University of Miami***	
2013	Dartmouth College	
2014	*None*	
2015	Wake Forest University	
2016	State University of New York At Buffalo*	Population Health Collaborative, Kaleida Health, Erie County Medical Center, Roswell Park Comprehensive Cancer Center, HEALTHeLINK, and VA Western New York Healthcare System

Notes:

1. Institutions with asterisks (*) are Carnegie R1 universities
2. Institutions with bold names are institutions that are ultimately included in the treatment group
3. Otherwise stated, CTSA partnership started in the year the CTSA hub started receiving the CTSA award

APPENDIX B. Networks measures with x-axis as normalized years

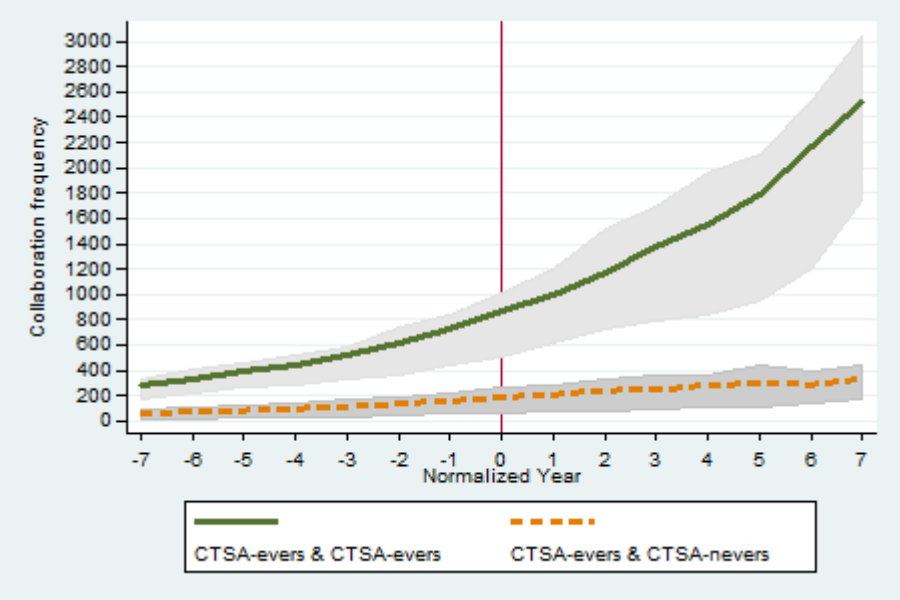


Figure B1. Collaboration frequency with CTSA-evers

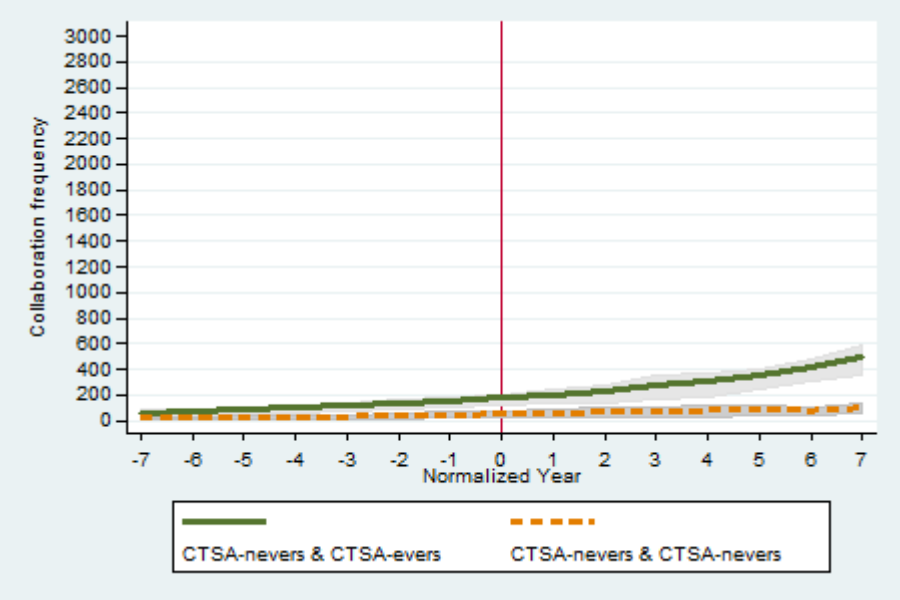


Figure B2. Collaboration frequency with CTSA-nevers

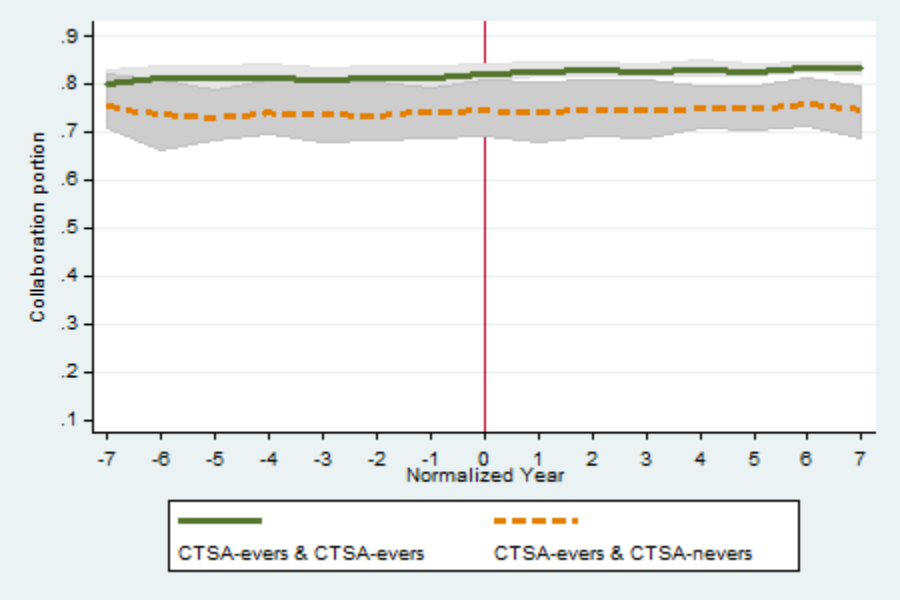


Figure B3. Collaboration portion with CTSA-evers

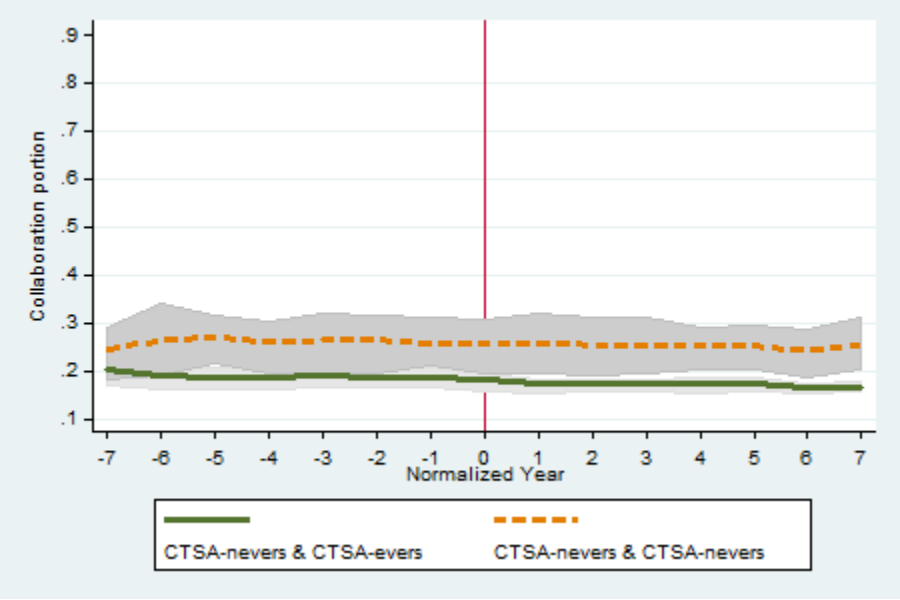


Figure B4. Collaboration portion with CTSA-nevers

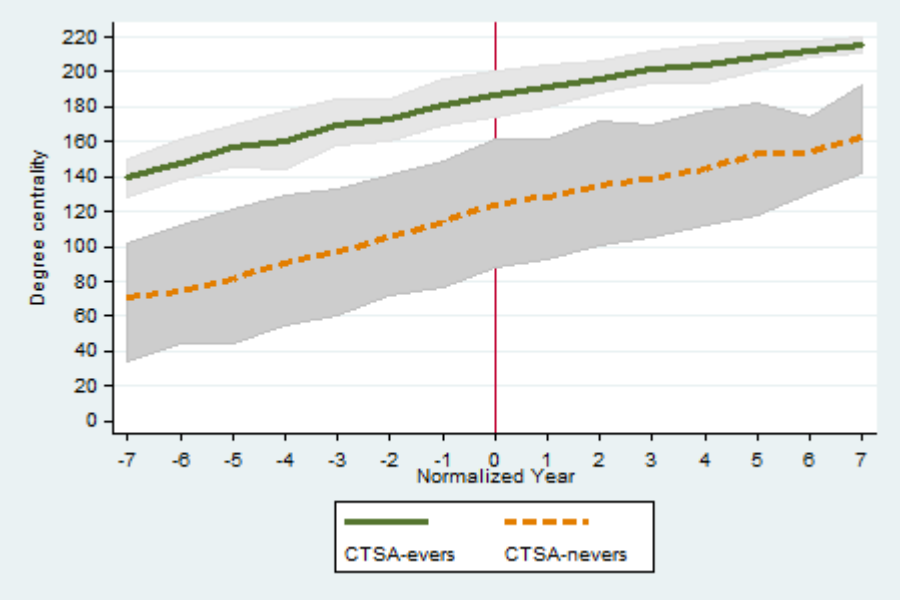


Figure B5. Degree centrality trend of institutions in CTSA-ever group and CTSA-never group

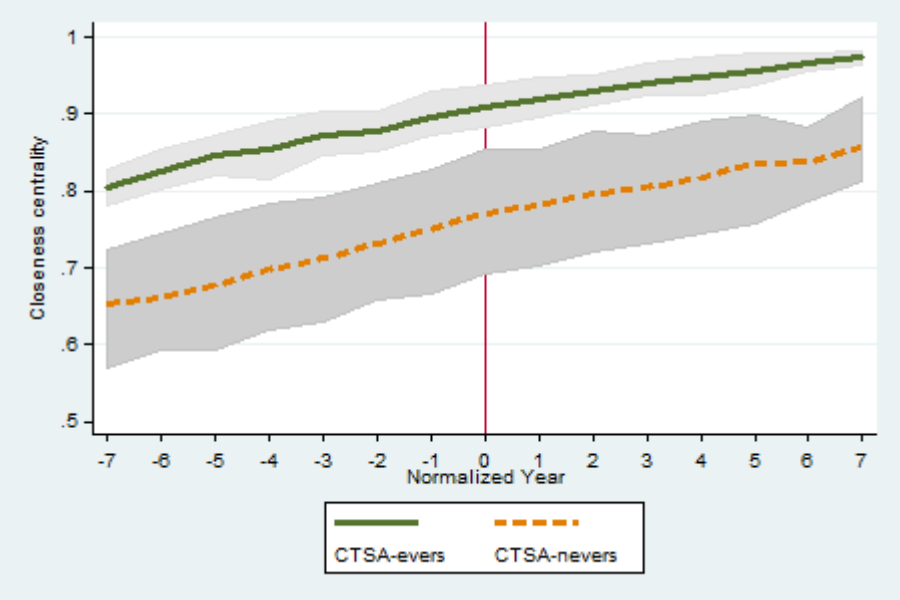


Figure B6. Closeness centrality trend of institutions in CTSA-ever group and CTSA-never group

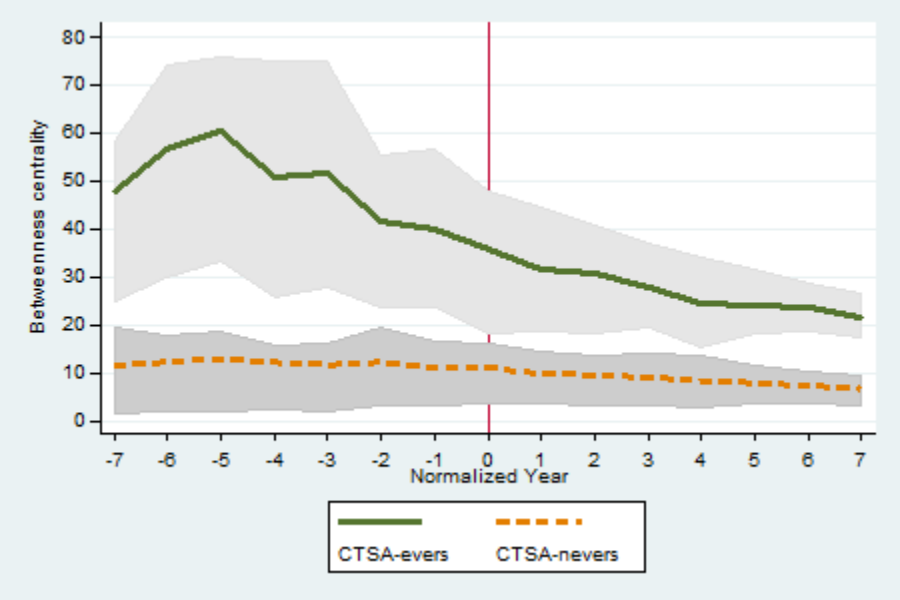


Figure B7. Betweenness centrality trend of institutions in CTSA-ever group and CTSA-never group

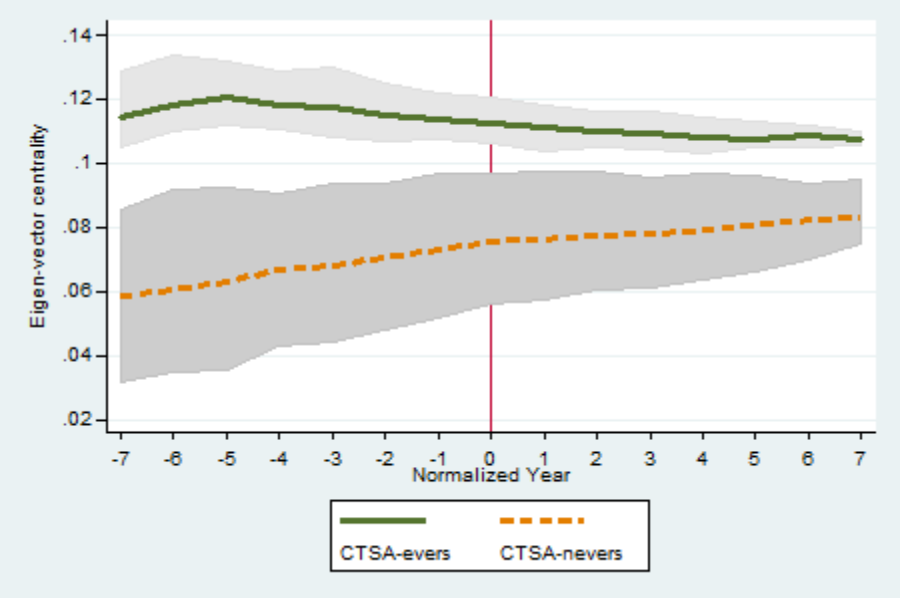
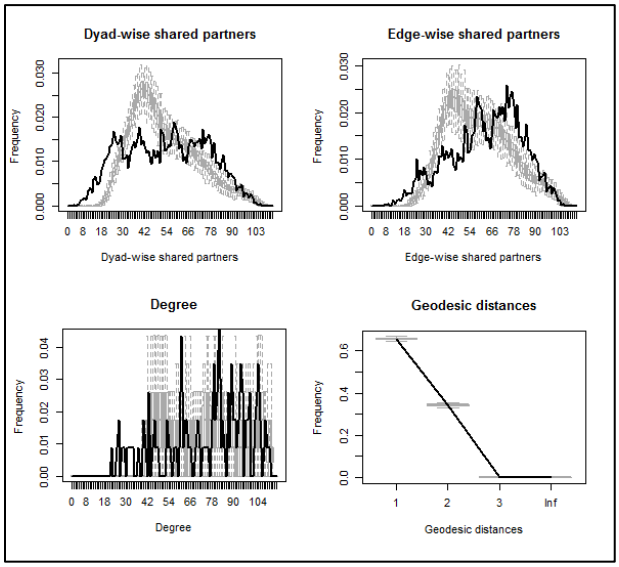


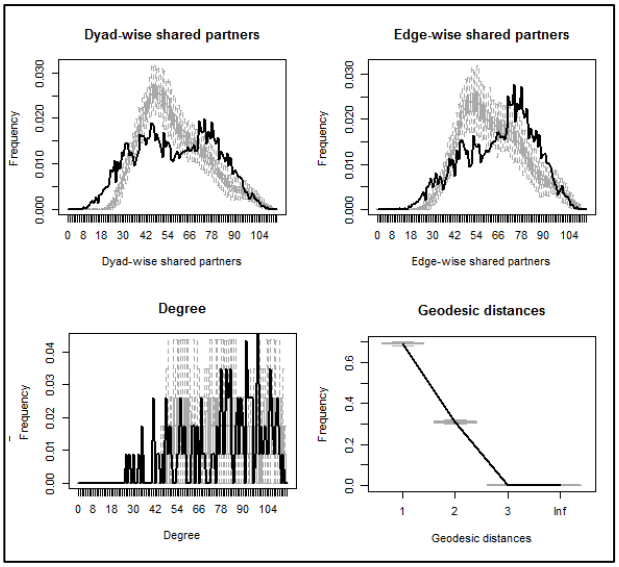
Figure B8. Eigen-vector centrality trend of institutions in CTSA-ever group and CTSA-never group

APPENDIX C. Goodness-of-fit assessment of binary ERGMs

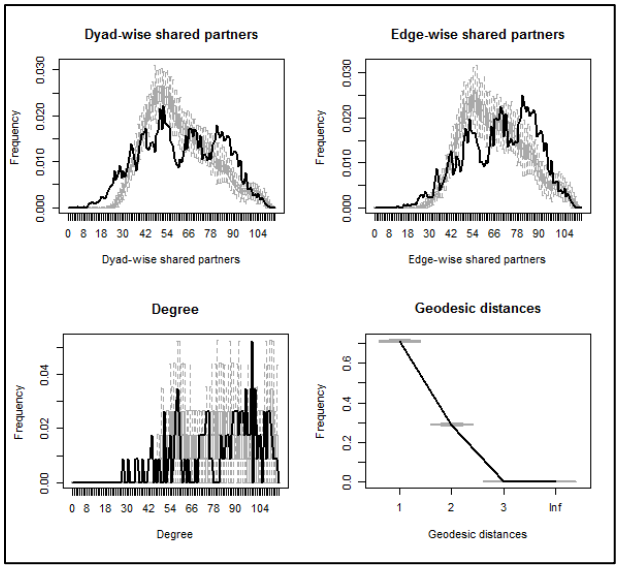
2009



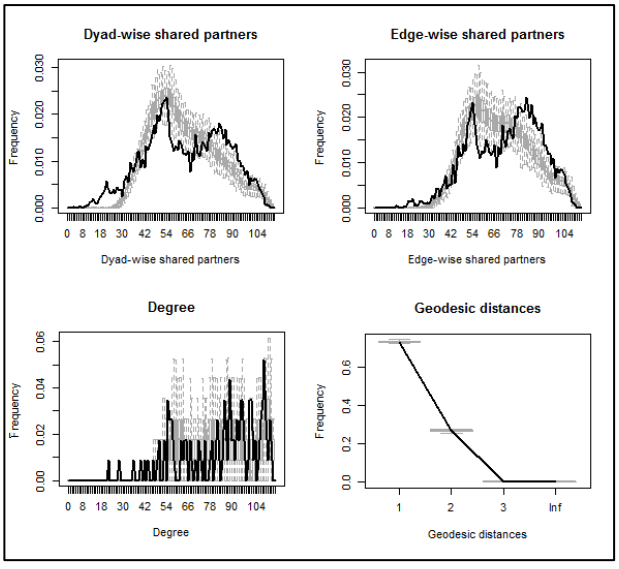
2010



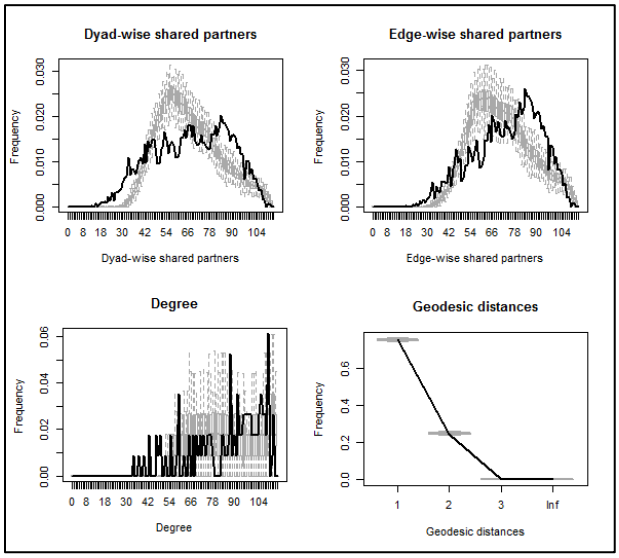
2011



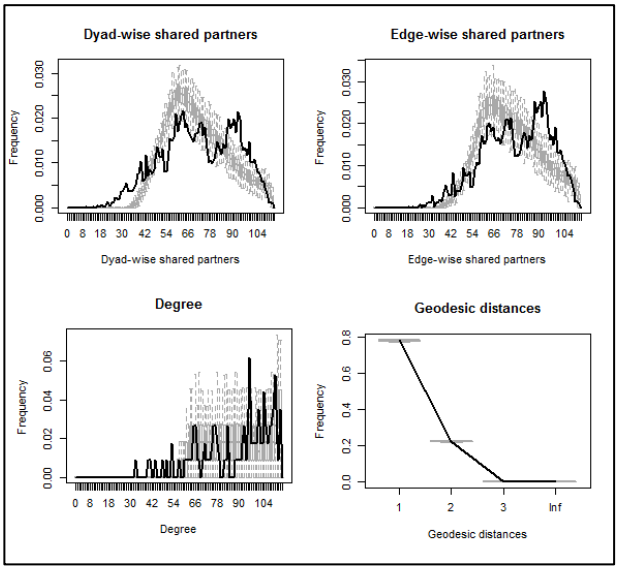
2012



2013



2014



2015

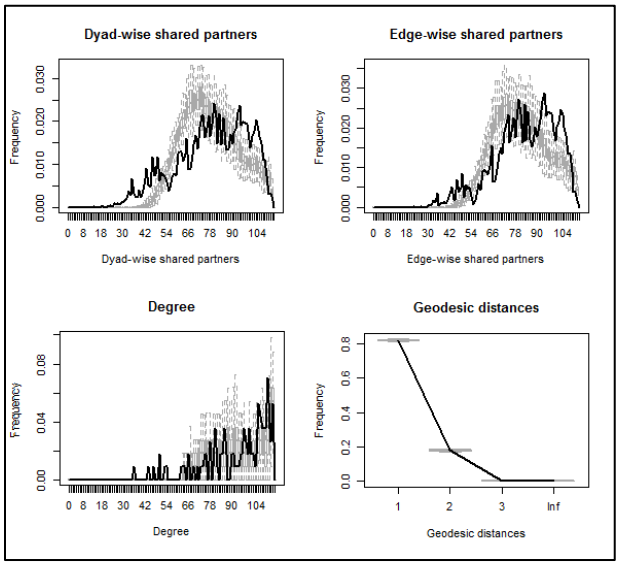


Figure C1. Goodness-of-fit assessment of binary ERGM from 2009 to 2015

APPENDIX D. Classification of basic science, clinical science, multidisciplinary science, and other science based on NSF's classification of fields of study

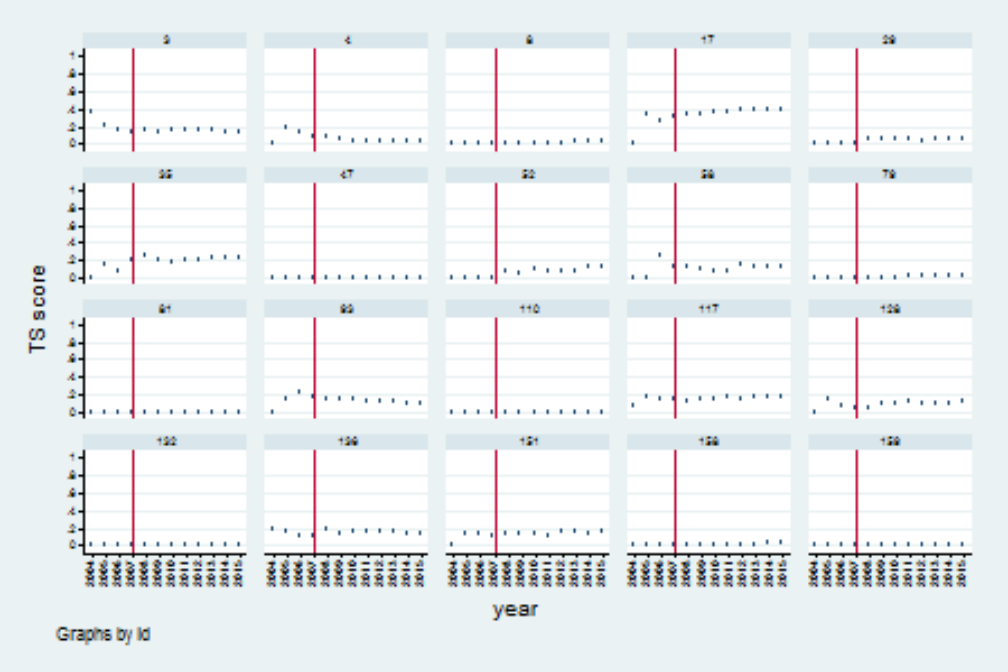
#	Counts	Web of Science Category	Non-Clinical	Clinical	Multi-disc	Non-S&E	Name on NSF list
1	115402	Biochemistry & Molecular Biology	X				Biochemistry
2	86350	Neurosciences	X				Neuroscience
3	65035	Cell Biology	X				Cell Biology and Anatomy
4	55905	Multidisciplinary Sciences			X		NA
5	47710	Oncology		X			Oncology
6	41626	Immunology	X				Immunology
7	37356	Genetics & Heredity	X				Genetics, General
8	37293	Public, Environmental & Occupational Health		X			Occupational Health and Industrial Hygiene
9	31937	Psychiatry		X			Psychiatry
10	30352	Pharmacology & Pharmacy		X			Pharmacy (PharmD [United States] PharmD, BS/BPharm [Canada])
11	29829	Endocrinology & Metabolism		X			Endocrinology and Metabolism
12	28471	Physiology	X				Physiology, General
13	27199	Clinical Neurology		X			Neurology
14	23593	Biochemical Research Methods	X				Biochemistry
15	23257	Medicine, Research & Experimental		X			Medicine (MD)
16	23140	Radiology, Nuclear Medicine & Medical Imaging		X			Nuclear Radiology
17	22414	Chemistry, Multidisciplinary	X				Chemistry, Other
18	22399	Microbiology	X				Microbiology, General
19	21071	Cardiac & Cardiovascular Systems	X				Cardiovascular Science
20	19993	Biophysics	X				Biophysics
21	19358	Infectious Diseases		X			Infectious Disease
22	18593	Biotechnology & Applied Microbiology	X				Microbiology, General
23	18444	Hematology		X			Hematology
24	18320	Peripheral Vascular Disease		X			Vascular Surgery
25	17960	Virology	X				Virology
26	15271	Surgery		X			General Surgery
27	14767	Pediatrics		X			Pediatrics
28	13974	Psychology, Clinical	X				Clinical Psychology
29	13493	Psychology	X				Psychology, General
30	13300	Engineering, Biomedical	X				Biomedical/Medical Engineering
31	12892	Developmental Biology	X				Developmental Biology and Embryology
32	12744	Ophthalmology		X			Ophthalmology
33	12189	Medicine, General & Internal		X			Internal Medicine
34	11732	Substance Abuse				X	Substance Abuse/Addiction Counseling
35	11469	Psychology, Developmental	X				Developmental and Child Psychology
36	11310	Mathematical & Computational Biology	X				Biomathematics and Bioinformatics, Other
37	11305	Respiratory System		X			Pulmonary Disease
38	11018	Chemistry, Organic	X				Organic Chemistry
39	10869	Toxicology	X				Toxicology
40	10744	Urology & Nephrology		X			Urology
41	10456	Gastroenterology & Hepatology		X			Gastroenterology
42	10263	Chemistry, Medicinal	X				Medicinal and Pharmaceutical Chemistry
43	10149	Behavioral Sciences				X	Behavioral Sciences
44	10104	Nutrition & Dietetics		X			Dietetics and Clinical Nutrition Services, Other
45	9858	Psychology, Experimental	X				Experimental Psychology
46	9498	Biology	X				Biology/Biological Sciences, General
47	9337	Geriatrics & Gerontology				X	Gerontology
48	9138	Health Care Sciences & Services		X			Health/Health Care Administration/Management
49	8887	Chemistry, Analytical	X				Analytical Chemistry
50	8838	Chemistry, Physical	X				Chemical Physics
51	8782	Obstetrics & Gynecology		X			Obstetrics and Gynecology
52	7722	Psychology, Multidisciplinary	X				Psychology, Other
53	7363	Statistics & Probability	X				Mathematical Statistics and Probability
54	7175	Health Policy & Services				X	Public Policy Analysis
55	6912	Gerontology				X	Gerontology
56	6483	Pathology		X			Pathology
57	6209	Computer Science, Interdisciplinary Applications	X				Computer Science, Other
58	5984	Critical Care Medicine		X			Critical Care Medicine
59	5934	Environmental Sciences	X				Environmental Science
60	5919	Nanoscience & Nanotechnology	X				NA
61	5645	Neuroimaging		X			Diagnostic Radiology
62	5611	Sport Sciences		X			Sports Medicine

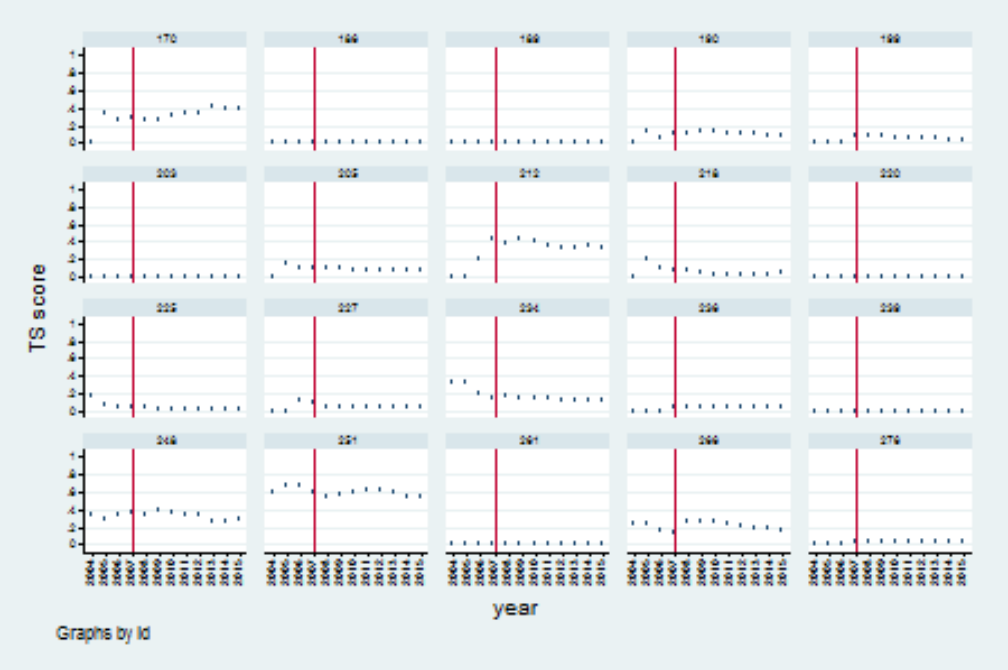
#	Counts	Web of Science Category	Non-Clinical	Clinical	Multi-disc	Non-S&E	Name on NSF list
63	5415	Rehabilitation		X			Physical and Rehabilitation Medicine
64	5325	Materials Science, Biomaterials	X				Materials Science
65	5250	Materials Science, Multidisciplinary	X				Materials Science
66	5207	Parasitology	X				Parasitology
67	5043	Transplantation		X			Nephrology
68	5027	Audiology & Speech-Language Pathology		X			Audiology/Audiologist and Speech-Language Pathology/Pathologist
69	4761	Social Sciences, Biomedical				X	Social Sciences, Other
70	4664	Nursing		X			Nursing—Registered Nurse Training (RN, ASN, BSN, MSN)
71	4422	Zoology	X				Zoology/Animal Biology
72	4383	Orthopedics		X			Orthopedics/Orthopedic Surgery
73	4302	Otorhinolaryngology		X			Otolaryngology
74	4268	Rheumatology		X			Rheumatology
75	4069	Reproductive Biology	X				Reproductive Biology
76	3981	Cell & Tissue Engineering	X				Biomedical/Medical Engineering
77	3972	Optics	X				Optics/Optical Sciences
78	3720	Medical Informatics				X	Medical Informatics
79	3671	Dentistry, Oral Surgery & Medicine		X			Dentistry (DDS, DMD
80	3261	Evolutionary Biology	X				Evolutionary Biology
81	3152	Spectroscopy	X				Radiation Biology/Radiobiology
82	3049	Family Studies				X	Human Development and Family Studies, General
83	3023	Acoustics	X				Acoustics
84	2992	Anesthesiology		X			Anesthesiology
85	2977	Psychology, Biological	X				Psychology, Other
86	2909	Physics, Atomic, Molecular & Chemical	X				Atomic/Molecular Physics
87	2884	Dermatology		X			Dermatology
88	2866	Plant Sciences	X				Plant Sciences, General
89	2758	Physics, Applied	X				Physics, Other
90	2579	Veterinary Sciences		X			Veterinary Sciences/Veterinary Clinical Sciences, General (Cert, MS, PhD)
91	2523	Linguistics				X	Linguistics
92	2391	Tropical Medicine		X			Community Health and Preventive Medicine
93	2304	Psychology, Social	X				Social Psychology
94	2211	Engineering, Electrical & Electronic	X				Electrical, Electronics, and Communications Engineering
95	2118	Allergy		X			Allergies and Immunology
96	2059	Anatomy & Morphology	X				Anatomy
97	1940	Ecology	X				Ecology
98	1937	Chemistry, Inorganic & Nuclear	X				Inorganic Chemistry
99	1740	Physics, Condensed Matter	X				Solid State and Low-Temperature Physics
100	1683	Social Sciences, Interdisciplinary				X	Social Sciences, Other
101	1633	Polymer Science	X				Polymer Chemistry
102	1607	Sociology				X	Sociology
103	1576	Social Work				X	Social Work
104	1561	Food Science & Technology	X				Food Science
105	1522	Emergency Medicine		X			Emergency Medicine
106	1467	Computer Science, Information Systems	X				Computer and Information Sciences, General
107	1368	Education & Educational Research				X	Education, General
108	1360	Psychology, Educational				X	Educational Psychology
109	1287	Education, Scientific Disciplines				X	Science Teacher Education/General Science Teacher Education
110	1220	Entomology	X				Entomology
111	1212	Computer Science, Artificial Intelligence	X				Artificial Intelligence and Robotics
112	1183	Medical Laboratory Technology		X			Laboratory Medicine
113	1180	Demography				X	Demography and Population Studies
114	1153	Information Science & Library Science				X	Library Science/Librarianship
115	1151	Women's Studies				X	Women’s Studies
116	1128	Mycology	X				Mycology
117	1116	Crystallography	X				Atomic/Molecular Physics
118	1116	Engineering, Environmental	X				Environmental/Environmental Health Engineering
119	1030	Chemistry, Applied	X				Chemistry, Other
120	1024	Economics				X	Economics, General
121	886	Imaging Science & Photographic Technology		X			Diagnostic Radiology
122	859	Psychology, Applied	X				Psychology, Other
123	835	Mathematics, Interdisciplinary Applications	X				Applied Mathematics, Other
124	827	Education, Special				X	Special Education and Teaching, General
125	766	Integrative & Complementary Medicine		X			Alternative and Complementary Medicine and Medical Systems, Other

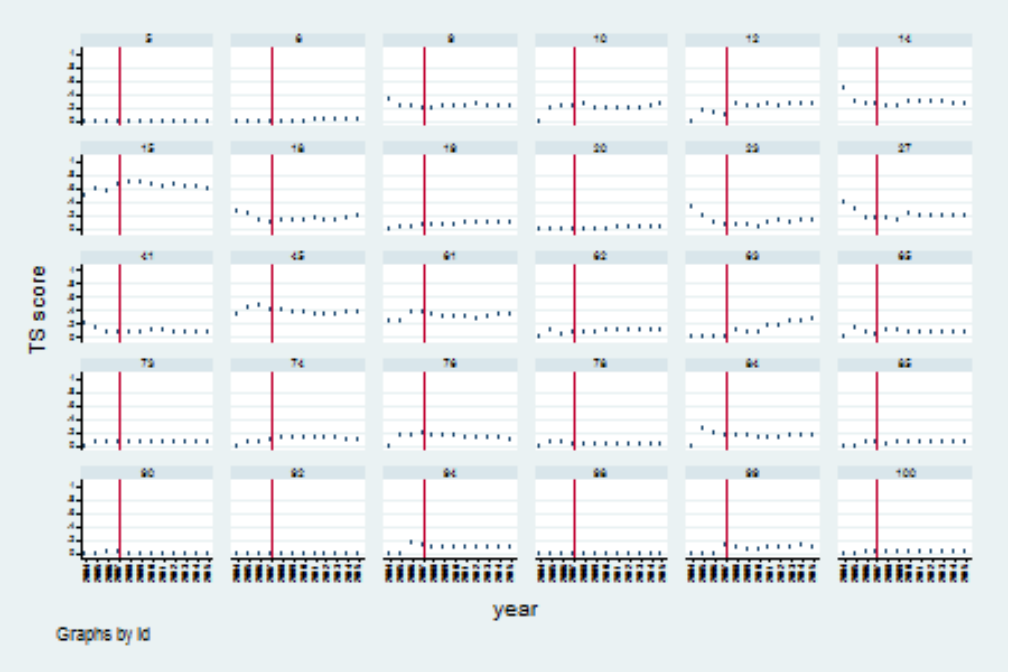
#	Counts	Web of Science Category	Non-Clinical	Clinical	Multi-disc	Non-S&E	Name on NSF list
126	761	Physics, Mathematical	X				Theoretical and Mathematical Physics
127	751	Instruments & Instrumentation	X				Instrumentation Technology/Technician
128	667	Criminology & Penology				X	Criminology
129	657	Physics, Multidisciplinary	X				Physics, Other
130	648	Anthropology				X	Anthropology
131	638	Physics, Fluids & Plasmas	X				Plasma and High-Temperature Physics
132	614	Psychology, Mathematical	X				Psychology, Other
133	546	Communication				X	Communication Studies/Speech Communication and Rhetoric
134	533	Nuclear Science & Technology	X				Nuclear Engineering
135	512	Primary Health Care		X			Health/Health Care Administration/Management
136	505	Ethics				X	Ethics
137	477	Microscopy	X				Instrumentation Technology/Technician
138	470	Electrochemistry	X				Chemistry, Other
139	430	Marine & Freshwater Biology	X				Marine Biology and Biological Oceanography
140	418	Medical Ethics				X	Bioethics/Medical Ethics
141	379	Ergonomics		X			NA
142	376	Mathematics, Applied	X				Applied Mathematics
143	375	Social Sciences, Mathematical Methods				X	Social Sciences, Other
144	337	Agriculture, Multidisciplinary	X				Agriculture, Agriculture Operations, and Related Sciences, Other
145	336	Computer Science, Theory & Methods	X				Computer Science
146	309	Law				X	Law (LLB, JD)
147	307	Social Issues				X	Social Sciences, Other
148	281	Computer Science, Software Engineering	X				Computer Software Engineering
149	281	Ethnic Studies				X	Ethnic, Cultural Minority, and Gender Studies, Other
150	270	Engineering, Chemical	X				Chemical Engineering
151	254	Andrology		X			Urology
152	244	Medicine, Legal		X			Medical Scientist (MS, Ph.D.)
153	244	Meteorology & Atmospheric Sciences	X				Atmospheric Sciences and Meteorology, General
154	240	Water Resources	X				Hydrology and Water Resources Science
155	225	Engineering, Industrial	X				Industrial Engineering
156	224	Language & Linguistics				X	Linguistics
157	211	Agriculture, Dairy & Animal Science	X				Dairy Science
158	206	Transportation				X	Transportation/Transportation Management
159	205	Fisheries	X				Fishing and Fisheries Sciences and Management
160	205	Mechanics	X				Engineering Mechanics
161	165	Environmental Studies				X	Environmental Studies
162	154	Automation & Control Systems	X				Electrical, Electronics, and Communications Engineering
163	152	Religion				X	Religion/Religious Studies
164	151	Political Science				X	Political Science and Government, General
165	146	Computer Science, Cybernetics	X				Computer Science, Other
166	145	Engineering, Multidisciplinary	X				Engineering, General
167	121	Engineering, Mechanical	X				Mechanical Engineering
168	121	Geography				X	Geography
169	103	Geosciences, Multidisciplinary	X				Geological and Earth Sciences/Geosciences, Other
170	97	Robotics	X				Artificial Intelligence and Robotics
171	95	Management				X	Business Administration and Management, General
172	94	Hospitality, Leisure, Sport & Tourism				X	Parks, Recreation, and Leisure Studies
173	82	Biodiversity Conservation	X				Ecology, Evolution, Systematics, and Population Biology, Other
174	77	History & Philosophy Of Science				X	History and Philosophy of Science and Technology
175	74	Planning & Development				X	Land Use Planning and Management/Development
176	73	Energy & Fuels	X				Petroleum Engineering
177	71	Urban Studies				X	Urban Studies/Affairs
178	68	Mathematics	X				Mathematics, General
179	67	Operations Research & Management Science				X	Operations Research
180	65	Industrial Relations & Labor				X	Labor and Industrial Relations
181	62	Physics, Nuclear	X				Nuclear Physics
182	59	Construction & Building Technology	X				Construction Engineering
183	57	Computer Science, Hardware & Architecture	X				Computer Hardware Engineering
184	52	Telecommunications	X				Computer Systems Networking and Telecommunications
185	49	Materials Science, Coatings & Films	X				Materials Science
186	48	Business				X	Business/Managerial Economics
187	48	Physics, Particles & Fields	X				Elementary Particle Physics
188	42	Agronomy	X				Soil Science and Agronomy, General
189	41	Engineering, Civil	X				Civil Engineering, General
190	41	Thermodynamics	X				Mechanical Engineering
191	40	Engineering, Manufacturing	X				Manufacturing Engineering
192	39	Business, Finance				X	Finance, General
193	38	Public Administration				X	Public Administration

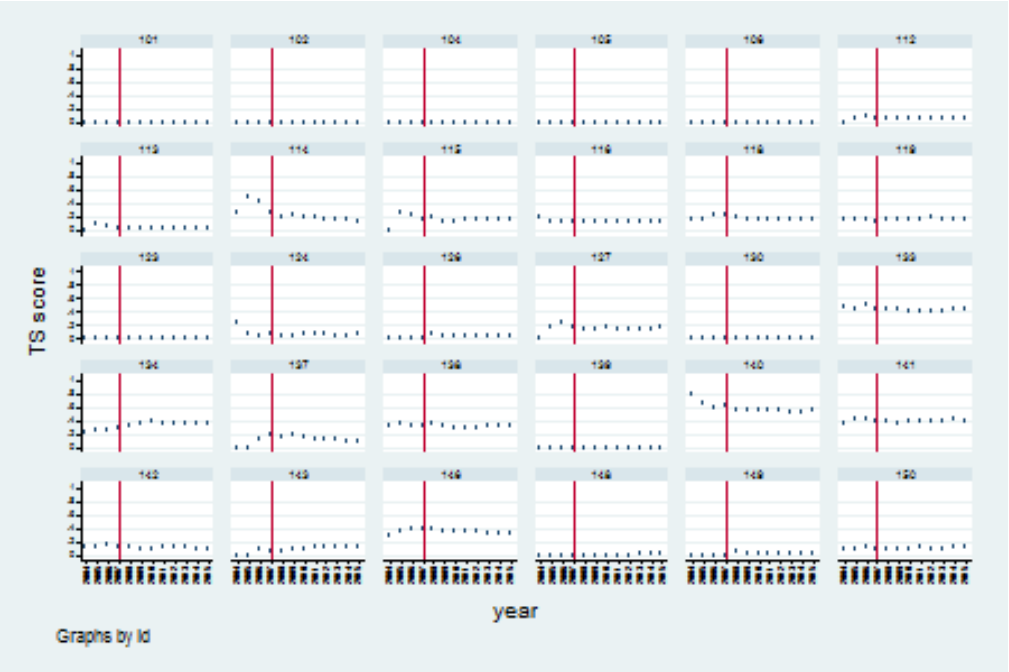
#	Counts	Web of Science Category	Non-Clinical	Clinical	Multi-disc	Non-S&E	Name on NSF list
194	35	History				X	History, General
195	35	Philosophy				X	Philosophy
196	33	Remote Sensing	X				Geological and Earth Sciences/Geosciences, Other
197	32	Astronomy & Astrophysics	X				Astronomy and Astrophysics, Other
198	31	Geography, Physical				X	Geography, Other
199	30	Metallurgy & Metallurgical Engineering	X				Metallurgical Engineering
200	28	Limnology	X				Aquatic Biology/Limnology
201	28	Psychology, Psychoanalysis	X				Psychoanalysis and Psychotherapy
202	27	Area Studies				X	Area Studies, Other
203	27	GREEN & SUSTAINABLE SCIENCE & TECHNOLOGY			X		
204	27	International Relations				X	International Relations and Affairs
205	25	History Of Social Sciences				X	History, Other
206	24	Ornithology	X				Animal Sciences, General
207	23	Geochemistry & Geophysics	X				Geochemistry
208	23	Horticulture	X				Agricultural and Horticultural Plant Breeding
209	23	Humanities, Multidisciplinary				X	Humanities/Humanistic Studies
210	19	Music				X	Music, General
211	18	Agricultural Engineering	X				Agricultural/Biological Engineering and Bioengineering
212	18	Soil Science	X				Soil Science and Agronomy, General
213	16	Transportation Science & Technology	X				Transportation and Highway Engineering
214	15	Materials Science, Ceramics	X				Materials Science
215	15	Oceanography	X				Oceanography, Chemical and Physical
216	14	Engineering, Aerospace	X				Aerospace, Aeronautical, and Astronautical Engineering
217	13	Materials Science, Characterization & Testing	X				Materials Science
218	11	Mineralogy	X				Geochemistry and Petrology
219	9	Film, Radio, Television				X	Radio and Television
220	9	Materials Science, Composites	X				Materials Science
221	9	Materials Science, Textiles	X				Materials Science
222	7	Cultural Studies				X	Ethnic, Cultural Minority, and Gender Studies, Other
223	5	Agricultural Economics & Policy				X	Agricultural Economics
224	5	Archaeology				X	Paleontology
225	5	Forestry	X				Forestry, General
226	5	Literature				X	Foreign Languages and Literatures, General
227	5	Mining & Mineral Processing	X				Mining and Mineral Engineering
228	2	Geology	X				Geology/Earth Science, General
229	2	Materials Science, Paper & Wood	X				Materials Science
230	1	Logic				X	Logic

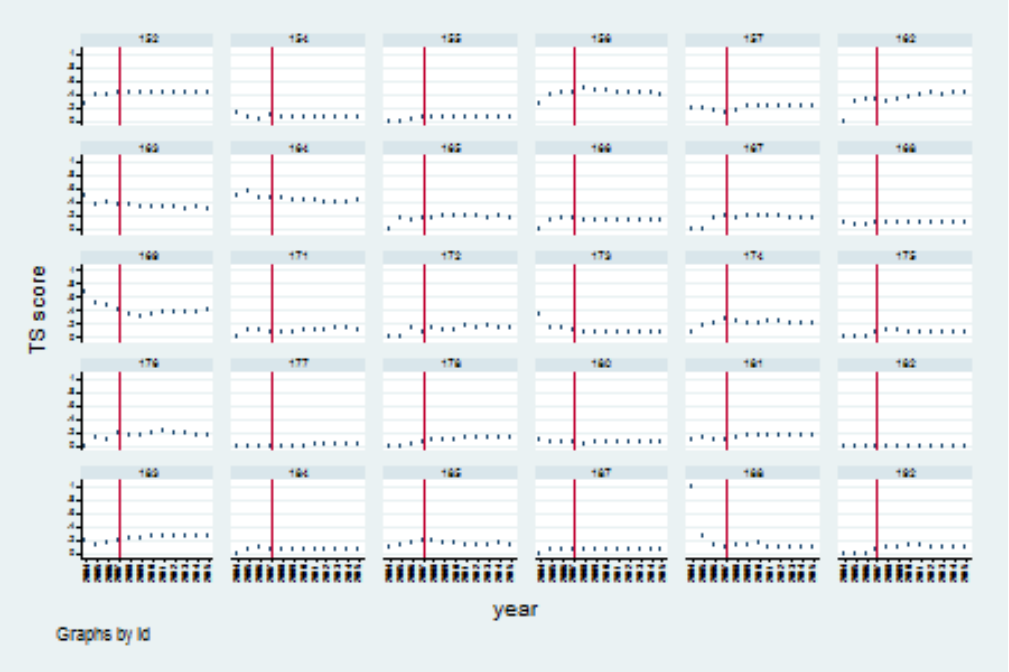
APPENDIX E. More examples the change in TS score by year for some publications

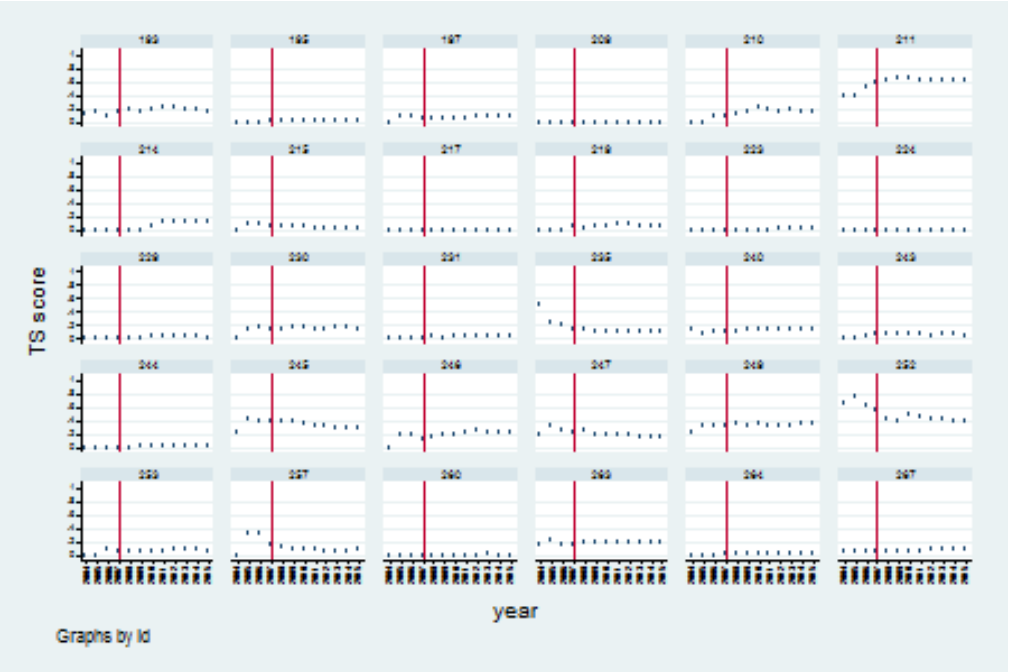


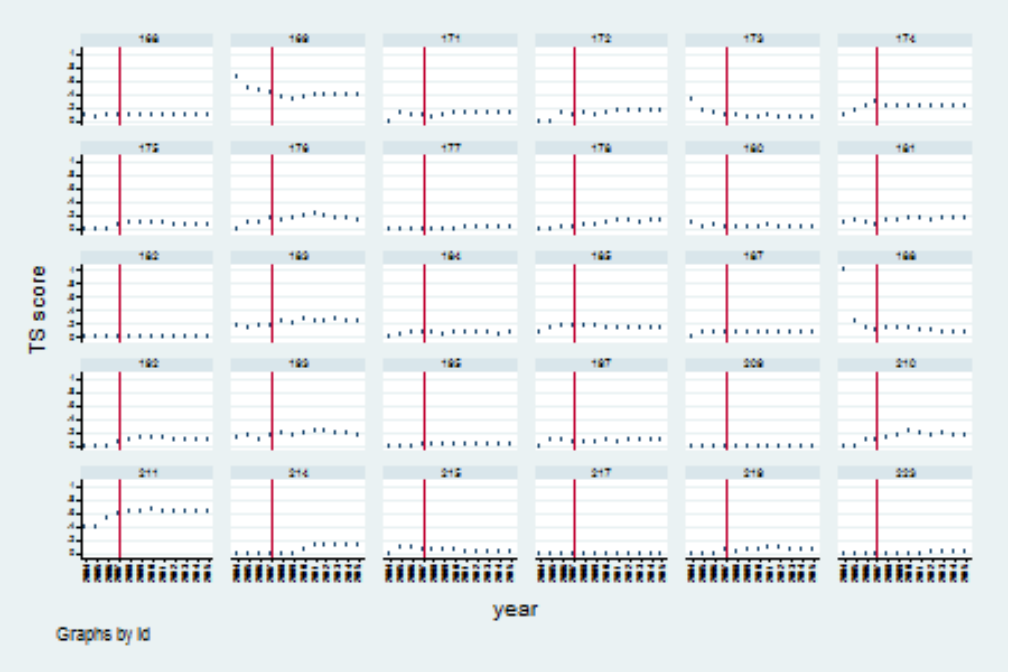


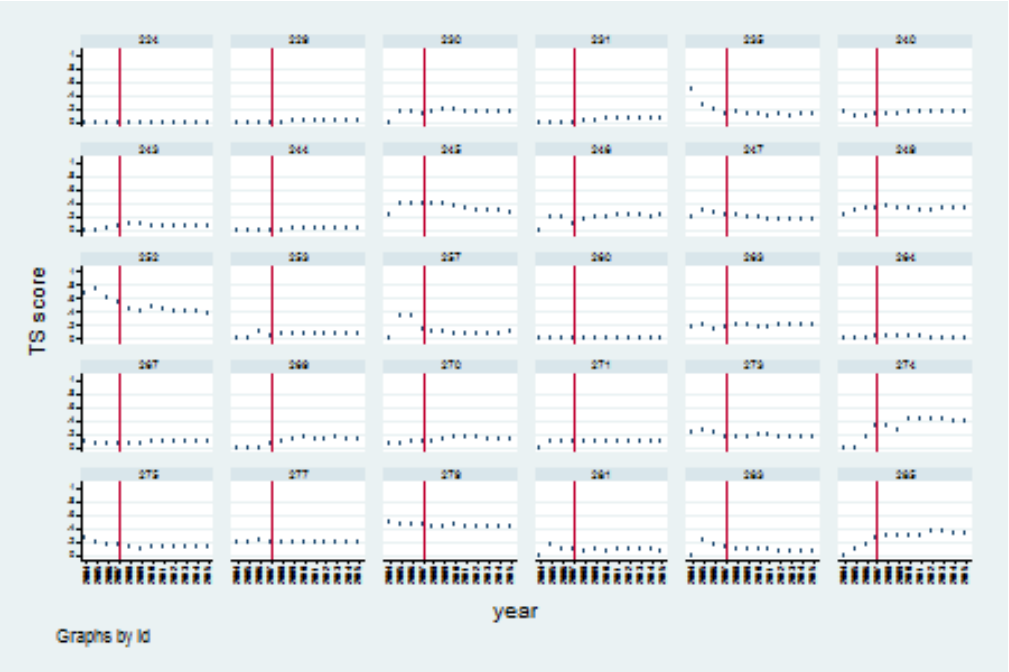

TS score
year
Graphs by id


TS score
year
Graphs by id


TS score
year
Graphs by id


TS score
year
Graphs by id


TS score
year
Graphs by Id


TS score
year
Graphs by Id

APPENDIX F. Institution matching based on propensity score of taking advantage of the CTSA award

Institution Code	Matched institution code	Institution	Propensity score	CTSA treatment dummy	CTSA cohort year
2	107	Boston Coll	0.0012	0	2010
100	107	Univ Texas Arlington	0.0076	0	2010
9	107	Clemson Univ	0.0117	0	2010
107	46	Univ Wisconsin Milwaukee	0.0183	1	2010
46	107	Syracuse Univ	0.0190	0	2010
27	107	Kansas State Univ	0.0217	0	2010
49	107	Texas Tech Univ	0.0280	0	2010
22	107	Georgia State Univ	0.0289	0	2010
4	107	Brandeis Univ	0.0317	0	2010
63	107	Univ Cent Florida	0.0324	0	2010
62	107	Univ Cal Santa Cruz	0.0374	0	2010
36	90	Oregon State Univ	0.0409	0	2008
59	90	Univ Cal Riverside	0.0463	0	2008
13	90	CUNY	0.0480	0	2008
16	90	Florida Int Univ	0.0539	0	2008
43	90	SUNY Albany	0.0556	0	2008
72	90	Univ Houston	0.0560	0	2008
90	72	Univ Notre Dame	0.0595	1	2008
6	90	CALTECH	0.0677	0	2008
18	90	George Mason Univ	0.0685	0	2008
1	90	Arizona State Univ	0.0695	0	2008
25	90	Iowa State Univ	0.0705	0	2008
89	90	Univ N Texas	0.0798	0	2008
110	90	Virginia Tech	0.0849	0	2008
40	7	Rice Univ	0.0859	0	2006
7	33	Carnegie Mellon Univ	0.1109	1	2006
33	7	Northeastern Univ	0.1204	0	2006
79	10	Univ Maryland College Park	0.1275	0	2008
38	10	Princeton Univ	0.1299	0	2008
111	10	Washington State Univ	0.1338	0	2008
10	111	Colorado State Univ	0.1340	1	2008
61	10	Univ Cal Santa Barbara	0.1356	0	2008
66	10	Univ Colorado Boulder	0.1443	0	2008
71	10	Univ Hawaii Manoa	0.1733	0	2008
41	21	Rutgers State Univ	0.1769	0	2007
32	21	N Carolina State Univ	0.1908	0	2007
96	21	Univ S Carolina	0.1939	0	2007
68	21	Univ Delaware	0.2074	0	2007
21	17	Georgia Inst Technol	0.2185	1	2007
17	21	Florida State Univ	0.2192	0	2007
101	21	Univ Texas Austin	0.2403	0	2007
92	21	Univ Oregon	0.2550	0	2007
47	21	Temple Univ	0.2580	0	2007
39	45	Purdue Univ	0.3224	1	2008
29	45	MIT	0.3462	1	2008
45	54	SUNY Stony Brook	0.3608	0	2009
54	45	Univ Arkansas	0.3691	1	2009
76	97	Univ Kansas	0.3798	1	2011
97	76	Univ S Florida	0.3858	0	2011
30	20	Michigan State Univ	0.4088	0	2010
114	20	W Virginia Univ	0.4226	0	2010
84	20	Univ Mississippi	0.4279	0	2010
44	20	SUNY Buffalo	0.4290	0	2010
78	20	Univ Louisville	0.4313	0	2010
20	78	Georgetown Univ	0.4315	1	2010

74	20	Univ Illinois Urbana Champaign	0.4789	0	2010
70	20	Univ Georgia	0.4826	0	2010
48	80	Texas A&M Univ	0.5046	0	2010
51	80	Tulane Univ	0.5343	0	2010
91	80	Univ Oklahoma	0.5366	0	2010
80	91	Univ Massachusetts Amherst	0.5369	1	2010
87	91	Univ New Mexico	0.5485	1	2010
86	81	Univ Nebraska	0.5950	0	2012
55	81	Univ Cal Berkeley	0.5990	0	2012
5	81	Brown Univ	0.6095	0	2012
81	102	Univ Miami	0.6313	1	2012
57	102	Univ Cal Irvine	0.6317	1	2010
102	57	Univ Texas Dallas	0.6389	0	2010
113	24	Wayne State Univ	0.6427	0	2008
24	113	Indiana Univ	0.6523	1	2008
28	98	Louisiana State Univ	0.6763	0	2010
98	99	Univ So Cal	0.6878	1	2010
73	99	Univ Illinois Chicago	0.6879	1	2009
19	99	George Washington Univ	0.6913	1	2010
99	95	Univ Tennessee	0.6965	0	2006
95	99	Univ Rochester	0.6985	1	2006
50	99	Tufts Univ	0.7130	1	2008
85	104	Univ Missouri	0.7311	1	2007
31	104	NYU	0.7349	1	2009
104	115	Univ Virginia	0.7549	0	2006
115	104	Yale Univ	0.7552	1	2006
11	53	Columbia Univ	0.7697	1	2006
8	53	Case Western Reserve Univ	0.7756	1	2007
53	8	Univ Arizona	0.7802	0	2007
67	103	Univ Connecticut	0.7915	0	2008
103	67	Univ Utah	0.7960	1	2008
64	67	Univ Chicago	0.8080	1	2007
65	67	Univ Cincinnati	0.8080	1	2009
3	67	Boston Univ	0.8114	1	2008
75	67	Univ Iowa	0.8243	1	2007
42	67	Stanford Univ	0.8260	1	2008
109	67	Virginia Commonwealth Univ	0.8274	1	2010
77	67	Univ Kentucky	0.8452	1	2011
34	67	Northwestern Univ	0.8541	1	2008
83	67	Univ Minnesota	0.8579	1	2011
35	67	Ohio State Univ	0.8762	1	2008
82	67	Univ Michigan	0.8771	1	2007
52	67	Univ Alabama	0.8799	1	2008
37	67	Penn State Univ	0.8838	1	2011
14	67	Duke Univ	0.8844	1	2006
15	67	Emory Univ	0.8900	1	2007
12	67	Cornell Univ	0.8931	1	2007
94	67	Univ Pittsburgh	0.8977	1	2006
26	67	Johns Hopkins Univ	0.9047	1	2007
60	67	Univ Cal San Diego	0.9050	1	2010
56	67	Univ Cal Davis	0.9051	1	2006
105	67	Univ Washington Seattle	0.9131	1	2007
112	67	Washington Univ St Louis	0.9187	1	2007
108	67	Vanderbilt Univ	0.9191	1	2007
93	67	Univ Penn	0.9272	1	2006
106	67	Univ Wisconsin Madison	0.9306	1	2006
23	67	Harvard Univ	0.9344	1	2008
69	67	Univ Florida	0.9366	1	2009
88	67	Univ N Carolina Chapel Hill	0.9415	1	2008

APPENDIX G. Data summary and correlation between variables having all Carnegie R1 universities in the sample

Table G1. Summary of data

Variable	Obs	Mean	Std.dev	Min	Max
Non-clinical science publications	1,725	215.61	220.83	1.00	1659.00
Mean forward citations (per paper)	1,725	49.76	23.12	8.88	220.41
Mean forward citations from clinical science (per paper)	1,725	6.30	3.88	0.00	30.15
Mean TS score (institution-year unit)	1,725	0.12	0.05	0.00	0.27
R&D (million $)	1,725	197.17	185.67	0.90	882.69
Researcher	1,725	1246.71	961.17	60.00	8511.00
Medical	1,725	0.58	0.49	0.00	1.00
Public	1,725	0.30	0.46	0.00	1.00
Degree centrality	1,725	141.28	54.58	4.00	228.00
Multidisciplinarity	1,725	2.01	0.14	1.29	2.49
Race diversity	1,725	1.60	0.26	0.50	2.38

* Note: Unit is institution-year

Table G2. Correlation between variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Non-clinical science publications	1.00										
(2) Mean forward citations (per paper)	0.16***	1.00									
(3) Mean forward citations from clinical science (per paper)	0.33***	0.61***	1.00								
(4) Mean TS score (institution-year unit)	0.26***	-0.10***	0.59***	1.00							
(5) R&D (million $)	0.82***	0.00	0.26***	0.35***	1.00						
(6) Researcher	0.76***	0..12***	0.34***	0.31***	0.71***	1.00					
(7) Medical	0.46***	0.01	0.43***	0.61***	0.54***	0.43***	1.00				
(8) Private	0.27***	0.20***	0.23***	0.08***	0.17***	0.04**	0.17***	1.00			
(9) Degree centrality	0.70***	-0.11***	0.20***	0.38***	0.73***	0.58***	0.57***	0.16***	1.00		
(10) Multidisciplinarity (institution-year unit)	0.15***	-0.48***	-0.17***	0.30***	0.20***	0.10***	0.16***	-0.08***	0.43***	1.00	
(11) Race diversity	0.03	-0.21***	0.03	0.27***	0.06**	0.22***	0.17***	0.13***	0.20***	0.33***	1.00

Note: *** p < 0.01, ** p< 0.05, * < 0.10

APPENDIX H. More DID regression results for test on Hypothesis 1

Table H1. Difference-in-difference OLS regression (not dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	31.40***	31.31***	30.64***	30.57***	13.19***	13.61***	11.69***	12.07***
CTSA	94.81***		95.76***		34.51***		37.19***	
POST	4.37**	4.37**	-11.35***	-11.46***	-5.15**	-4.84**	-2.46	-2.63
Researcher					23.71***	20.64***	21.99***	18.59***
R&D					0.19***	0.19***	0.20***	0.19***
Med school					0.93	-5.89	-2.43	-9.05
Private					25.32**		24.87**	
Multidisciplinarity					23.71***	27.86***	22.89**	22.33**
Centrality					0.00	-0.01	-0.00	-0.05
Constant	26.42**	71.78***	13.10	58.90***	-58.87***	-33.07*	-55.63***	-16.18
Observation	1380	1380	1380	1380	1380	1380	1380	1380
σ_u	67.92	83.51	67.96	84.05	41.18	53.70	41.20	56.31
σ_e	19.54	19.54	17.68	17.68	15.84	15.84	15.08	15.08
ρ	0.92	0.95	0.94	0.96	0.87	0.92	0.88	0.93
Year FE	NO	NO	YES	YES	NO	NO	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Table H2. Difference-in-difference Poisson regression (not dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	0.01***	0.10***	0.08***	0.08***	0.09***	0.09***	0.05*	0.05*
CTSA	1.52***		1.55***		1.08***		1.10***	
POST	0.16***	0.16***	-0.07***	-0.07***	-0.10***	-0.10***	-0.05*	-0.05*
Researcher					0.07***	0.06***	0.05***	0.05***
R&D					0.00***	0.00***	0.00***	0.00***
Med school					0.27***	0.13*	0.20**	0.06
Private					0.01		0.02	
Multidisciplinarity					0.79***	0.88***	0.99***	1.00***
Centrality					0.00***	0.00***	0.00***	0.00***
Constant	3.27***		3.06***		1.02***		0.68**	
Observation	1380	1380	1380	1380	1380	1380	1380	1380
Pseudo R^2	0.013	0.125	0.047	0.238	0.104	0.214	0.104	0.253
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Table H3. Difference-in-difference Poisson regression results (dropping CTSA direct supported publications, not dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	0.04*	0.04*	0.02	0.02	0.03	0.04	-0.01	-0.01
CTSA	1.51***		1.53***		1.11***		1.10***	
POST	0.16***	0.16***	-0.03	-0.03	-0.08***	-0.08***	-0.01	-0.01
Researcher					0.07***	0.07***	0.06***	0.05***
R&D					0.00*	0.00*	0.00***	0.00***
Med school					-0.27***	0.13	0.19**	0.05
Private					0.00		0.02	
Multidisciplinarity					0.76***	0.85***	1.06***	1.07***
Centrality					0.00**	0.00***	0.00***	0.00***
Constant	3.27***		3.07***		1.14***		0.57*	
Observation	1380	1380	1380	1380	1380	1380	1380	1380
Pseudo R^2	0.013	0.079	0.042	0.182	0.107	0.156	0.103	0.198
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Table H4. Difference-in-difference OLS regression results (dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	14.01***	13.94***	13.71***	13.65***	8.99***	9.15***	8.63***	8.70***
CTSA	50.35***		50.66***		25.91***		25.64***	
POST	4.35***	4.36***	-3.72*	-3.77**	-4.33***	-3.76***	-1.58	-1.69
Researcher					8.11***	6.71***	7.29***	5.50**
R&D					0.16***	0.14***	0.17***	0.15***
Med school					4.05	-4.79	1.39	-6.67*
Private					17.62***		17.73***	
Multidisciplinarity					12.53**	16.92***	15.69**	15.09**
Centrality					0.05**	0.04*	0.08***	0.04
Constant	26.72***	44.07***	19.14***	36.59***	-27.53**	-13.72	-34.54***	-9.09
Observation	1080	1080	1080	1080	1080	1080	1080	1080
σ_u	36.98	44.38	37.01	44.56	19.66	33.16	19.68	33.94
σ_e	11.55	11.55	10.38	10.38	10.02	10.02	9.51	9.51
ρ	0.91	0.94	0.93	0.95	0.79	0.92	0.81	0.93
Year FE	NO	NO	YES	YES	NO	NO	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Table H5. Difference-in-difference Poisson regression results (dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	0.06**	0.06**	0.04*	0.04*	0.07***	0.07***	0.05*	0.05*
CTSA	1.06***		1.07***		0.82***		0.81***	
POST	0.16***	0.16***	-0.03	-0.03	-0.10***	-0.09***	-0.04	-0.04
Researcher					0.09***	0.08***	0.05*	0.04
R&D						0.00	0.00***	0.00*
Med school					0.26***	0.15*	0.19**	0.07
Private							0.00	
Multidisciplinarity					0.78***	0.87***	1.00***	1.01***
Centrality					0.00***	0.00***	0.00***	0.00***
Constant	3.28***		3.09***		1.05***		0.58*	
Observation	1080	1080	1080	1080	1080	1080	1080	1080
Pseudo R^2	0.007	0.057	0.019	0.137	0.106	0.121	0.074	0.153
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

Table H6. Difference-in-difference Poisson regression results (dropping CTSA direct supported publications, dropping unmatched institutions)

	Without covariates				With covariates			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
CTSA*POST	-0.01	-0.01	-0.03	-0.03	0.02	0.02	-0.00	-0.00
CTSA	1.05***		1.06***		0.84***		0.82***	
POST	0.16***	0.16***	0.01	0.01	-0.09***	-0.08***	-0.02	-0.02
Researcher					0.10***	0.09***	0.06**	0.05*
R&D					-0.00	-0.00*	0.00	0.00
Med school					0.27***	0.16*	0.19**	0.08
Private					0.00		0.00	
Multidisciplinarity					0.79***	0.88***	1.06***	1.07***
Centrality					0.00***	0.00***	0.00***	0.00***
Constant	3.28***		3.10***		1.06***		0.48	
Observation	1080	1080	1080	1080	1080	1080	1080	1080
Pseudo R^2	0.007	0.033	0.016	0.109	0.112	0.091	0.077	0.127
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Institution FE	NO	YES	NO	YES	NO	YES	NO	YES

APPENDIX I. Counts and portion of CTSA supported papers by institution-year

Table I1. Counts of all CTSA supported publications regardless of forward citation counts

Institution	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Sum
Boston Univ			1	4	12	27	29	33	42	49	197
Case Western Reserve Univ	3	29	45	91	121	187	346	374	370	313	1879
Columbia Univ	1	7	28	51	70	100	140	204	236	264	1101
Cornell Univ			14	41	55	86	110	173	124	150	753
Duke Univ		5	26	34	50	45	54	46	60	92	412
Emory Univ		14	124	147	187	220	267	210	208	207	1584
George Washington Univ						4	16	33	37	62	152
Georgetown Univ			2	7	11	25	51	38	54	64	252
Harvard Univ				27	98	203	230	271	280	314	1423
Johns Hopkins Univ			18	56	133	155	178	204	413	387	1544
NYU				1	11	36	55	78	94	75	350
Northwestern Univ			45	84	128	182	199	182	81	141	1042
Ohio State Univ		1	36	95	143	190	182	148	138	118	1051
Penn State Univ			1	1	4	8	40	38	60	59	211
Stanford Univ				1	13	29	64	92	102	135	436
Tufts Univ			1	10	17	41	36	44	51	65	265
Univ Alabama			2	9	24	42	64	83	85	80	389
Univ Arkansas				1	8	34	74	73	114	93	397
Univ Cal Davis		4	18	30	54	55	63	60	52	67	403
Univ Cal Irvine			1			12	24	34	42	119	232
Univ Cal Los Angeles	1		1		2	52	249	358	327	335	1325
Univ Cal San Diego					1	12	38	54	59	75	239
Univ Chicago			2	18	40	48	67	62	479	319	1035
Univ Cincinnati		2	5	39	85	101	248	414	422	531	1847
Univ Florida		4	12	15	19	45	423	105	95	150	868
Univ Illinois Chicago					14	44	53	54	75	85	325
Univ Iowa			20	35	55	52	59	74	57	62	414
Univ Kansas				1		1	31	37	62	54	186
Univ Kentucky		2	5	5	4	13	69	84	114	127	423
Univ Massachusetts Amherst						34	14	69	57	57	231
Univ Miami						2	1	23	23	20	69
Univ Michigan	1		10	34	89	131	134	175	145	174	893
Univ Minnesota				2	3	8	44	79	178	172	486
Univ New Mexico			1	3	14	16	85	124	97	109	449
Univ N Carolina Chapel Hill			7	39	81	108	131	177	200	225	968
Univ Penn		2	38	52	70	115	113	143	153	180	866
Univ Pittsburgh	3	190	202	173	301	405	411	504	434	460	3083
Univ Rochester		1	33	35	57	69	134	85	128	142	684
Univ So Cal					2	10	72	95	58	78	315
Univ Utah			2	13	22	38	68	73	98	131	445
Univ Washington Seattle			22	49	90	304	171	146	161	164	1107
Univ Wisconsin Madison			16	44	146	169	197	204	292	275	1343
Vanderbilt Univ		1	10	57	85	167	245	297	347	367	1576
Virginia Commonwealth Univ						30	57	84	71	77	319
Washington Univ St Louis			20	104	151	232	285	335	354	359	1840
Yale Univ	3	8	22	55	81	122	131	161	393	265	1241

Table I2. Counts of CTSA supported non-clinical publications with five or more forward citations

Institution	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Sum
Boston Univ				1	1	4	3	5	8	10	41
Case Western Reserve Univ	1	8	12	28	23	40	79	79	80	48	424
Columbia Univ			1	6	12	13	33	42	41	36	229
Cornell Univ			2	9	2	13	14	16	17	17	100
Duke Univ			3	1	2	3	1	2	1	2	15
Emory Univ		3	23	28	33	38	31	34	42	23	279
George Washington Univ							4	5	5	16	36
Georgetown Univ				1	4	4	7	5	6	5	50
Harvard Univ				2	15	32	25	33	25	26	202
Johns Hopkins Univ				7	26	32	28	40	64	43	291
NYU					2	8	13	15	16	9	85
Northwestern Univ			15	25	26	39	40	27	14	27	248
Ohio State Univ			7	13	30	30	36	26	14	15	185
Penn State Univ						3	6	9	8	9	48
Stanford Univ					2	6	10	9	12	20	78
Tufts Univ				4	4	3	4	4	4	11	39
Univ Alabama			1	2	7	7	7	16	9	10	69
Univ Arkansas					3	7	15	16	28	16	101
Univ Cal Davis			6	5	5	6	12	10	14	9	74
Univ Cal Irvine						2	5	5	4	18	50
Univ Cal Los Angeles						9	45	58	54	49	295
Univ Cal San Diego					1	3	5	8	8	12	50
Univ Chicago				5	9	8	12	17	118	63	254
Univ Cincinnati		1		4	7	10	30	40	56	58	251
Univ Florida			1		1	4	99	14	12	25	194
Univ Illinois Chicago					2	11	12	9	17	18	83
Univ Iowa			6	8	16	13	17	12	13	10	102
Univ Kansas							7	5	4	3	23
Univ Kentucky			1				13	15	27	28	103
Univ Massachusetts Amherst						6	1	16	18	10	68
Univ Miami								5	3	3	13
Univ Michigan			3	4	6	12	28	29	27	27	163
Univ Minnesota				2	1	4	7	13	14	21	89
Univ New Mexico				1	1	1	12	18	3	14	53
Univ N Carolina Chapel Hill			1	5	6	7	15	28	24	19	140
Univ Penn			6	11	8	18	15	20	24	22	145
Univ Pittsburgh		20	28	28	46	67	62	63	67	71	463
Univ Rochester			6	9	8	6	26	17	23	15	116
Univ So Cal							6	12	5	7	33
Univ Utah			1	2	6	7	12	8	7	16	68
Univ Washington Seattle			2	5	12	34	24	15	12	18	136
Univ Wisconsin Madison			1	5	23	23	20	23	32	32	191
Vanderbilt Univ			2	9	11	40	44	58	54	63	323
Virginia Commonwealth Univ						5	3	11	12	10	53
Washington Univ St Louis			4	12	9	26	45	38	39	44	270
Yale Univ		2	3	7	15	25	16	24	85	47	420

Table I3. Portion of CTSA supported publications with five or more forward citations among all publications

Institution	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Boston Univ				0	0	0.01	0.01	0.01	0.02	0.03
Case Western Reserve Univ	0	0.02	0.03	0.05	0.04	0.09	0.17	0.18	0.19	0.13
Columbia Univ			0	0.01	0.02	0.02	0.05	0.06	0.06	0.06
Cornell Univ			0	0.02	0.01	0.02	0.02	0.03	0.03	0.03
Duke Univ			0.01	0	0.01	0.01	0	0.01	0	0.01
Emory Univ		0.01	0.06	0.06	0.07	0.07	0.06	0.07	0.08	0.05
George Washington Univ						0.01	0.05	0.06	0.04	0.13
Georgetown Univ				0.01	0.03	0.02	0.04	0.04	0.06	0.05
Harvard Univ				0	0.01	0.02	0.01	0.02	0.02	0.02
Johns Hopkins Univ				0.01	0.04	0.04	0.03	0.05	0.08	0.06
NYU					0	0.02	0.03	0.04	0.04	0.02
Northwestern Univ			0.04	0.05	0.05	0.07	0.08	0.05	0.02	0.05
Ohio State Univ			0.02	0.03	0.07	0.08	0.09	0.06	0.04	0.05
Penn State Univ						0.01	0.02	0.03	0.04	0.03
Stanford Univ					0	0.01	0.01	0.01	0.01	0.03
Tufts Univ				0.02	0.02	0.02	0.02	0.02	0.02	0.05
Univ Alabama			0	0.01	0.02	0.02	0.02	0.04	0.02	0.04
Univ Arkansas					0.03	0.06	0.14	0.16	0.32	0.17
Univ Cal Davis			0.02	0.02	0.01	0.01	0.03	0.02	0.03	0.02
Univ Cal Irvine						0	0.01	0.01	0.01	0.06
Univ Cal Los Angeles						0.01	0.05	0.08	0.08	0.07
Univ Cal San Diego					0	0	0.01	0.01	0.01	0.01
Univ Chicago				0.01	0.02	0.02	0.03	0.03	0.24	0.20
Univ Cincinnati		0	0	0.01	0.02	0.03	0.10	0.13	0.20	0.22
Univ Florida			0	0	0.01	0.01	0.27	0.05	0.04	0.08
Univ Illinois Chicago					0.01	0.04	0.04	0.03	0.07	0.07
Univ Iowa			0.02	0.02	0.04	0.03	0.05	0.03	0.03	0.03
Univ Kansas				0	0	0	0.03	0.02	0.02	0.02
Univ Kentucky			0	0	0	0	0.05	0.07	0.14	0.16
Univ Massachusetts Amherst						0.02	0	0.04	0.04	0.03
Univ Miami								0.02	0.02	0.02
Univ Michigan			0	0	0.01	0.01	0.03	0.03	0.03	0.03
Univ Minnesota				0	0	0.01	0.01	0.02	0.03	0.03
Univ New Mexico				0.01	0.01	0.01	0.09	0.19	0.03	0.12
Univ N Carolina Chapel Hill			0	0.01	0.01	0.01	0.02	0.03	0.03	0.03
Univ Penn			0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03
Univ Pittsburgh		0.04	0.04	0.04	0.07	0.09	0.09	0.09	0.09	0.10
Univ Rochester			0.02	0.03	0.03	0.02	0.09	0.05	0.07	0.07
Univ So Cal					0		0.01	0.03	0.01	0.02
Univ Utah			0	0.01	0.01	0.02	0.03	0.02	0.03	0.05
Univ Washington Seattle			0	0.01	0.01	0.04	0.03	0.02	0.02	0.03
Univ Wisconsin Madison			0	0.01	0.04	0.03	0.03	0.04	0.05	0.05
Vanderbilt Univ			0	0.02	0.02	0.07	0.06	0.08	0.09	0.11
Virginia Commonwealth Univ						0.03	0.02	0.07	0.07	0.06
Washington Univ St Louis			0.01	0.02	0.01	0.04	0.06	0.05	0.06	0.07
Yale Univ		0	0	0.01	0.02	0.04	0.02	0.03	0.12	0.07

Table I4. Portion of CTSA supported non-clinical publications with five or more forward citations among all non-clinical science publications

Institution	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Boston Univ				0.00	0.00	0.01	0.01	0.01	0.02	0.03
Case Western Reserve Univ	0.00	0.02	0.03	0.05	0.04	0.08	0.15	0.16	0.17	0.11
Columbia Univ			0.00	0.01	0.02	0.02	0.05	0.06	0.06	0.05
Cornell Univ			0.00	0.02	0.00	0.02	0.02	0.02	0.03	0.03
Duke Univ			0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Emory Univ		0.01	0.05	0.06	0.07	0.07	0.06	0.06	0.07	0.04
George Washington Univ							0.04	0.05	0.04	0.13
Georgetown Univ				0.01	0.03	0.02	0.04	0.03	0.04	0.04
Harvard Univ				0.00	0.01	0.02	0.01	0.02	0.01	0.02
Johns Hopkins Univ				0.01	0.03	0.04	0.03	0.05	0.07	0.05
NYU					0.00	0.02	0.03	0.03	0.04	0.02
Northwestern Univ			0.03	0.05	0.05	0.07	0.07	0.05	0.02	0.04
Ohio State Univ			0.02	0.03	0.07	0.07	0.08	0.06	0.03	0.04
Penn State Univ						0.01	0.02	0.02	0.02	0.02
Stanford Univ					0.00	0.01	0.01	0.01	0.01	0.02
Tufts Univ				0.02	0.02	0.01	0.02	0.02	0.02	0.05
Univ Alabama			0.00	0.01	0.02	0.02	0.02	0.04	0.02	0.03
Univ Arkansas					0.03	0.06	0.13	0.14	0.25	0.16
Univ Cal Davis			0.02	0.01	0.01	0.01	0.03	0.02	0.03	0.02
Univ Cal Irvine						0.00	0.01	0.01	0.01	0.05
Univ Cal Los Angeles						0.01	0.05	0.07	0.07	0.06
Univ Cal San Diego					0.00	0.00	0.00	0.01	0.01	0.01
Univ Chicago				0.01	0.02	0.01	0.02	0.03	0.21	0.17
Univ Cincinnati		0.00		0.01	0.02	0.03	0.09	0.11	0.16	0.16
Univ Florida			0.00		0.00	0.01	0.25	0.04	0.03	0.06
Univ Illinois Chicago					0.01	0.03	0.04	0.03	0.06	0.06
Univ Iowa			0.02	0.02	0.04	0.03	0.04	0.02	0.03	0.02
Univ Kansas							0.02	0.02	0.02	0.01
Univ Kentucky			0.00				0.05	0.06	0.12	0.12
Univ Massachusetts Amherst						0.02	0.00	0.04	0.04	0.02
Univ Miami								0.02	0.02	0.01
Univ Michigan			0.00	0.00	0.01	0.01	0.03	0.03	0.03	0.03
Univ Minnesota				0.00	0.00	0.01	0.01	0.02	0.02	0.03
Univ New Mexico				0.01	0.01	0.01	0.09	0.15	0.03	0.10
Univ N Carolina Chapel Hill			0.00	0.01	0.01	0.01	0.02	0.03	0.03	0.02
Univ Penn			0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02
Univ Pittsburgh		0.03	0.04	0.04	0.07	0.08	0.09	0.08	0.08	0.08
Univ Rochester			0.02	0.03	0.02	0.02	0.07	0.05	0.07	0.05
Univ So Cal							0.01	0.03	0.01	0.02
Univ Utah			0.00	0.01	0.01	0.02	0.03	0.02	0.02	0.04
Univ Washington Seattle			0.00	0.01	0.01	0.04	0.03	0.01	0.01	0.02
Univ Wisconsin Madison			0.00	0.01	0.04	0.03	0.03	0.03	0.04	0.05
Vanderbilt Univ			0.00	0.02	0.02	0.06	0.06	0.08	0.07	0.09
Virginia Commonwealth Univ						0.03	0.02	0.06	0.07	0.04
Washington Univ St Louis			0.01	0.02	0.01	0.04	0.06	0.05	0.05	0.06
Yale Univ		0.00	0.00	0.01	0.02	0.03	0.02	0.03	0.11	0.06

APPENDIX J. Model estimation results for testing mediation effects

Table J1. Detailed result of mediation tests of inter-organizational collaboration

Year	Condition 1		Condition 2		Condition 3		Condition 4		Direct effect	Portion of indirect effect
	Coef c	YES/NO	Coef a	YES/NO	|c| -| c'|	YES/NO	Indirect effect	YES/NO		
2009	.0762***	YES	-.01036	NO	0.1309	YES	-.0002	NO	.0764	NA
2010	.1121***	YES	-.0536**	YES	0.0193	YES	-.0011	NO	.1132	NA
2011	.3141***	YES	.01660	NO	0.1669	YES	.0003	NO	.3138**	NA
2012	.1922***	YES	.01583	NO	NA	NO	.0003	NO	.1919	NA
2013	.3950***	YES	.0521**	YES	0.1742	YES	.0010	NO	.3940***	NA
2014	.6004***	YES	.0682***	YES	0.2084	YES	.0010*	YES	.5993***	.0017
2015	.6374***	YES	.0933***	YES	0.2917	YES	.0014**	YES	.6359***	.0022

Table J2. Detailed result of mediation tests of Shannon diversity

Year	Condition 1		Condition 2		Condition 3		Condition 4		Direct effect	Portion of indirect effect
	Coef c	YES/NO	Coef a	YES/NO	|c| -| c'|	YES/NO	Indirect effect	YES/NO		
2009	.0762***	YES	0.0020	NO	0.0018	YES	-0.0016	NO	0.0778	NA
2010	.1121***	YES	-0.0029	NO	0.0015	YES	0.0011	NO	0.1110	NA
2011	.3141***	YES	0.0024	NO	NA	NO	0.0003	NO	.3139**	NA
2012	.1922***	YES	-0.0135	NO	-0.0018	NO	-0.002	NO	0.1939	NA
2013	.3950***	YES	0.0001	NO	NA	NO	0.0003	NO	0.1919	NA
2014	.6004***	YES	0.0007	NO	-0.0004	NO	-0.0003	NO	.6006***	NA
2015	.6374***	YES	-0.0001	NO	-0.0022	NO	0.0000	NO	.6373***	NA

REFERENCES

Aboelela, S. W., Larson, E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S. A., ... & Gebbie, K. M. (2007). Defining interdisciplinary research: Conclusions from a critical *review of the literature. Health Services Research, 42*(1p1), 329-346.

Alberta, H. B., Cheng, A., Jackson, E. L., Pjecha, M., & Levine, A. D. (2015). Assessing state stem cell programs in the United States: how has state funding affected publication trends? *Cell Stem Cell*, *16*(2), 115-118.

Bagnol, D., Mansour, A., Akil, H., & Watson, S. J. (1997). Cellular localization and distribution of the cloned mu and kappa opioid receptors in rat gastrointestinal tract. *Neuroscience, 81*(2), 579-591.

Balconi, M., Breschi, S., & Lissoni, F. (2004). Networks of inventors and the role of academia: an exploration of Italian patent data. *Research Policy, 33*(1), 127-145.

Barron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Baumwol, K., Mortimer, S. T., Huerta, T. R., Norman, C. D., & Buchan, A. M. (2011). Promoting interdisciplinarity in the life sciences: a case study. *Research Evaluation, 20*(4), 283-292.

Bayer AG. (n.d.), *Small and large molecules*. Retrieved from http://pharma.bayer.com/en/innovation-partnering/technologies-and-trends/small-and-large-molecules/

Bhat, A., Kolibaba, K., Oda, T., Ohno-Jones, S., Heaney, C., & Druker, B. J. (1997). Interactions of CBL with BCR-ABL and CRKL in BCR-ABL-transformed myeloid cells. *Journal of Biological Chemistry, 272*(26), 16170-16175.

Bian, J., Xie, M., Topaloglu, U., Hudson, T., Eswaran, H., & Hogan, W. (2014). Social network analysis of biomedical research collaboration networks in a CTSA institution. *Journal of biomedical informatics, 52*, 130-140.

Blümel, C. (2017). Translational research in the science policy debate: a comparative analysis of documents. *Science and Public Policy, 45*(1), 24-35.

Borenstein, D. (2011, May 2), *Helping New Drugs Out of Research's 'Valley of Death'*. Retrieved from https://opinionator.blogs.nytimes.com/2011/05/02/helping-new-drugs-out-of-academias-valley-of-death/

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research policy*, *33*(4), 599-616.

Brint, S. (2005). Creating the future: 'New directions' in American research universities. *Minerva, 43*(1), 23-50.

Brown, B., Ortiz, C., & Dubé, K. (2018). Assessment of the Right-to-Try law: the pros and the cons. *Journal of Nuclear Medicine, 59*(10), 1492-1493.

Butler, D. (2008, June 11). *Translational research: Crossing the valley of death.* Retrieved from http://www.nature.com/news/2008/080611/full/453840a.html

Califf, R. M., & Ostroff, S. (2015). FDA as a catalyst for translation. *Science Translational Medicine, 7*(9), 1-4

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. *Handbook of Research on Teaching, 5,* 171-246.

Carpenter, C. R., Cone, D. C., & Sarli, C. C. (2014). Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine, 21*(10), 1160-1172.

Ciensinski, S. (2015, July 29) *Brazil Visits SRI to Discuss Its Economic Development Roadmap*. Retrieved from https://www.sri.com/blog/brazil-visits-sri-discuss-its-economic-development-roadmap

Clarivate Analytics (2018), *Web of Science, Trust the difference*. Retrieved from http://wokinfo.com/products_tools/multidisciplinary/?utm_source=false&utm_medium=false&utm_campaign=false&utm_source=false&utm_medium=false&utm_campaign=false

Clinical and Translational Science Awards Consortium (n.d.-a*). About the CTSA Consortium*. Retrieved from https://ctsacentral.org/about-us/ctsa/

Clinical and Translational Science Awards Consortium (n.d.-b.). *The Trial Innovation Network*, Retrieved from https://ctsacentral.org/about-us/trial-innovation-network/

Cockburn, I. M., & Henderson, R. M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *The Journal of Industrial Economics, 46*(2), 157-182.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Contopoulos-Ioannidis, D. G., Alexiou, G. A., Gouvias, T. C., & Ioannidis, J. P. (2008). Life cycle of translational research for medical interventions. *Science, 321*(5894), 1298-1299.

Corredoira, R. A., Goldfarb, B. D., & Shi, Y. (2018). Federal funding and the rate and direction of inventive activity. *Research Policy, 47*(9), 1777-1800.

Crommelin, D. J., Storm, G., Verrijk, R., de Leede, L., Jiskoot, W., & Hennink, W. E. (2003). Shifting paradigms: biopharmaceuticals versus low molecular weight drugs. *International Journal of Pharmaceutics, 266*(1-2), 3-16.

CTSA Principal Investigators. (2012). Preparedness of the CTSA's structural and scientific assets to support the mission of the National Center for Advancing Translational Sciences (NCATS). *Clinical and Translational Science, 5*(2), 121-129.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy, 38*(2), 293-305.

Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of Informetrics, 5*(1), 187-203.

Dodson, B. P., & Levine, A. D. (2015). Challenges in the translation and commercialization of cell therapies. *BMC Biotechnology, 15*(1), 70.

Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Research Policy, 11*(3), 147-162.

Drolet, B. C., & Lorenzi, N. M. (2011). Translational research: understanding the continuum from bench to bedside. *Translational Research*, 157(1), 1-5.

Dundar, H., & Lewis, D. R. (1998). Determinants of research productivity in higher education. *Research in Higher Education*, 39(6), 607-631.

EATRIS ERIC. (n.d.), *About EATRIS European Infrastructure for translational medicine*. Retrieved from http://eatris.eu/about/

Etzkowitz, H., Ranga, M., Benner, M., Guaranys, L., Maculan, A. M., & Kneller, R. (2008). Pathways to the entrepreneurial university: towards a global convergence. *Science and Public Policy*, *35*(9), 681-695.

European Commission. (2017, May 30). *News alert EATRIS: bridging the gap between medical research and clinical applications*. Retrieved from http://ec.europa.eu/research/index.cfm?pg=newsalert&year=2013&na=na-111113

Eynott, P. R., & Fages, C. (2014). Access Platform: A Streamlined Integrative Partnering Process at SANOFI to Commercialize University-Based Intellectual Property. In Chaguturu, R (Ed.), *Collaborative Innovation in Drug Discovery* (pp.319-340). Hoboken, NJ: Wiley.

Feldman, M. P., Lanahan, L., & Lendel, I. V. (2014). Experiments in the laboratories of democracy: State scientific capacity building. *Economic Development Quarterly, 28*(2), 107-131.

Fishburn, C. S. (2013). Translational research: the changing landscape of drug discovery. *Drug discovery today*, *18*(9), 487-494.

Fontelo, P., & Liu, F. (2011). Finding translational science publications in MEDLINE/PubMed with translational science filters. *Clinical and translational science*, *4*(6), 455-459.

Fries, R. C., Glave, S. A., & Radick, M. K. (2008, August). The benefits of Industry-University interaction. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1598-1601). IEEE.

Fukushima, M., & Kimura, Y. (n.d.) *Translational Research Informatics Center (TRI) takes the initiative to reduce disability through Japanese academic research*. Retrieved from http://s3-service-broker-live-19ea8b98-4d41-4cb4-be4c-d68f4963b7dd.s3.amazonaws.com/uploads/ckeditor/attachments/8326/TRi.pdf

Gabay, M. (2017). 21st century cures act. *Hospital pharmacy*, *52*(4), 264.

Garg, U. C., & Hassid, A. (1990). Nitric oxide-generating vasodilators inhibit mitogenesis and proliferation of BALBc 3T3 fibroblasts by a cyclic GMP-independent mechanism. *Biochemical and Biophysical Research Communications*, *171*(1), 474-479.

Gaughan, M., Melkers, J., & Welch, E. (2018). Differential Social Network Effects on Scholarly Productivity: An Intersectional Analysis. *Science, Technology, & Human Values*, *43*(3), 570-599.

Georgia Clinical & Translational Science Alliance (n.d.), *Georgia CTSA*. retrieved from http://georgiactsa.org/what-we-do/index.html

Gibbons, M., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Thousand Oaks, CA: SAGE Publications.

Gittelman, M. (2016). The revolution re-visited: Clinical and genetics research paradigms and the productivity paradox in drug discovery. *Research Policy*, 45(8), 1570-1585.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report, 8*(4), 597-606.

Grant, J., Cottrell, R., Cluzeau, F., & Fawcett, G. (2000). Evaluating "payback" on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ, 320*(7242), 1107-1111.

Han, X., Williams, S. R., & Zuckerman, B. L. (2018). A snapshot of translational research funded by the National Institutes of Health (NIH): A case study using behavioral and social science research awards and Clinical and Translational Science Awards funded publications. *PLoS One, 13*(5), e0196545.

Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside.

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library & information science research, 18*(4), 323-342.

He, Z. L., Geng, X. S., & Campbell-Hunt, C. (2009). Research collaboration and research output: A longitudinal study of 65 biomedical scientists in a New Zealand university. *Research Policy, 38*(2), 306-317.

Heinze, T., Shapira, P., Rogers, J. D., & Senker, J. M. (2009). Organizational and institutional influences on creativity in scientific research. *Research Policy, 38*(4), 610-623.

Hemlin, S., Allwood, C., & Martin, B. (2004). *Creative knowledge environments: The influences on creativity in research and innovation*. Northhampton, MA: Edward Elgar Pub.

Hicks, D., & Katz, J. S. (1996). Hospitals: the hidden research system. *Science and Public Policy*, *23*(5), 297-304.

Homans, G. (1950). *The human group.* New York, NY: Harcourt, Brace & World, Inc.

Huang, M. H., Yang, H. W., & Chen, D. Z. (2015). Increasing science and technology linkage in fuel cells: A cross citation analysis of papers and patents. *Journal of Informetrics*, *9*(2), 237-249.

Huerta, M. F., Farber, G. K., Wilder, E. L., Kleinman, D. V., Grady, P. A., Schwartz, D. A., & Tabak, L. A. (2005). NIH roadmap interdisciplinary research initiatives. *PLoS Computational Biology*, *1*(6), e59.

Huggett, B., Hodgson, J., & Lähteenmäki, R. (2011). Public biotech 2010 – the numbers. *Nature Biotechnology*, *29*, 585-591.

Hughes, M. E., Peeler, J., & Hogenesch, J. B. (2010). Network dynamics to evaluate performance of an academic institution. *Science Translational Medicine*, *2*(53), 53ps49.

Ihli, M. I. (2016). *Acknowledgement Lag and Impact: Domain Differences in Published Research Supported by the National Science Foundation* (master's thesis). Retrieved from https://trace.tennessee.edu/utk_gradthes/4045.

Indiana University Center for Postsecondary Research (n.d.). *The Carnegie Classification of Institutions of Higher Education, 2015 edition*, Bloomington, IN: Author.

International Committee of Medical Journal Editors (2019). *Defining the Role of Authors and Contributors*. Retrieved from http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html

Jabbehdari, S., & Walsh, J. P. (2017). Authorship norms and project structures in science. *Science, Technology, & Human Values, 42*(5), 872-900.

Japan External Trade Organizations (2017, May 8). *Japan's new approval system makes it easier for regenerative md3eical therapies to reach the market*. Retrieved from https://www.jetro.go.jp/usa/topics/japan-laws-regenerative-medicine-market.html

Kaiser, J. (2017, September 26). *Senate panel blocks NIH from revising translational research awards*. Retrieved from https://www.sciencemag.org/news/2017/09/senate-panel-blocks-nih-revising-translational-research-awards

Karmali, R. N., Jones, N. M., & Levine, A. D. (2010). Tracking and assessing the rise of state-funded stem cell research. *Nature Biotechnology, 28*(12), 1246.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*(1), 1-18.

Katz, N., Lazer, D., Arrow, H., & Contractor, N. (2004). Network theory and small groups. *Small group research, 35*(3), 307-332.

Kim, S. (2013), The concept of translational research and the requirements for the success (in Korean), *STEPI Insight, 15*(115), 1-25.

Kinch, M. S., & Hoyer, D. (2015). A history of drug development in four acts. *Drug Discovery Today, 20*(10), 1163-1168.

Knapke, J. M., Haynes, E. N., Kuhnell, P., & Tsevat, J. (2015). NIH grant awards as a metric of clinical and translational research training effectiveness. *Clinical and Translational Science, 8*(1), 52-56.

Knoepfler, P. S. (2015). From bench to FDA to bedside: US regulatory trends for new stem cell therapies. *Advanced Drug Delivery Reviews, 82*, 192-196.

Lander, B. (2013). Sectoral collaboration in biomedical research and development. *Scientometrics, 94*(1), 343-357.

Lander, B., & Atkinson-Grosjean, J. (2011). Translational science and the hidden research system in universities and academic hospitals: A case study. *Social Science & Medicine, 72*(4), 537-544.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673-702.

Leifeld, P., Cranmer, S. J., & Desmarais, B. A. (2018). Temporal exponential random graph models with btergm: Estimation and bootstrap confidence intervals. *Journal of Statistical Software, 83*(6), 1-36.

LENS.ORG. (n.d.). *About the Lens*. Retrieved from https://about.lens.org/

Li, D., Azoulay, P., & Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science, 356*(6333), 78-81.

Liu, G., Chen, G., Sinoway, L. I., & Berg, A. (2013). Assessing the impact of the NIH CTSA program on institutionally sponsored clinical trials. *Clinical and Translational Science, 6*(3), 196-200.

Leshner, A., Terry, S., Schultz, A., & Liverman, C. (2013). *The CTSA program at NIH: Opportunities for advancing clinical and translational research*. Washington, DC: National Academies Press.

Llewellyn, N., Carter, D. R., Rollins, L., & Nehl, E. J. (2018). Charting the Publication and Citation Impact of the NIH Clinical and Translational Science Awards (CTSA) Program From 2006 Through 2016. *Academic Medicine, 93*(8), 1162-1170.

Llewellyn, N.M., Carter, D. R., DiazGranados, D., Pelfrey, C., Rollins, L. & Nehl, E.J. (2019). Scope, Influence & Inter-disciplinary Collaboration: The Publication Portfolio of the NIH Clinical and Translational Science Awards (CTSA) Program from 2006 through 2017. *Evaluation & the Health Professions*. Manuscript Accepted for publication.

Lowe, P., & Phillipson, J. (2006). Reflexive interdisciplinary research: the making of a research programme on the rural economy and land use. *Journal of Agricultural Economics, 57*(2), 165-184.

Luke, D. A., Carothers, B. J., Dhand, A., Bell, R. A., Moreland-Russell, S., Sarli, C. C., & Evanoff, B. A. (2015). Breaking down silos: mapping growth of cross-disciplinary collaboration in a translational science initiative. *Clinical and Translational Science, 8*(2), 143-149.

Lynskey, M. J. (2006). The dismantling of redundant dichotomies: Biotechnology as an exemplar of university–industry collaboration. *Journal of Commercial Biotechnology, 12*(2), 127-147.

Martin, B. R., & Irvine, J. (1983). Assessing basic research: some partial indicators of scientific progress in radio astronomy. *Research Policy, 12*(2), 61-90.

Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research, 46*(5), 816-841.

McElroy, S. P., Jones, P. S., & Barrault, D. V. (2017). The SULSA Assay Development Fund: accelerating translation of new biology from academia to pharma. *Drug discovery today, 22*(2), 199-203.

McFarland, D. A., Moody, J., Diehl, D., Smith, J. A., & Thomas, R. J. (2014). Network ecology and adolescent social structure. *American sociological review, 79*(6), 1088-1121.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*(1), 415-444.

Medical Research Councils UK. (2017), *Translational research*. Retrieved from https://www.mrc.ac.uk/funding/science-areas/translation/

Meyer, M. (2000). Does science push technology? Patents citing scientific literature. *Research Policy, 29*(3), 409-434.

Moran, N. (2007). Public sector seeks to bridge 'valley of death'. *Nature Biotechnology, 25*(3), 266-267.

Morillo, F., Bordons, M., & Gómez, I. (2001). An approach to interdisciplinarity through bibliometric indicators. *Scientometrics, 51*(1), 203-222.

Morris, Z. S., Wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of the Royal Society of Medicine, 104*(12), 510-520.

Mowery, D. C. (1998). Collaborative R&D: How effective is it? *Issues in Science and Technology, 15*(1), 37-44.

Munos, B. H. (2014). Accelerating innovation in the bioscience revolution. In Chaguturu, R (Ed.), *Collaborative Innovation in Drug Discovery* (pp.195-212). Hoboken, NJ: Wiley.

Nam, S. (2015). *Globalization Strategy of Korean Bio Industry through US Bio-Venture Success Case Study* (in Korean). Seoul, Korea: Korean Institute of S&T Evaluation and Planning.

Narin, F., Pinski, G., & Gee, H. H. (1976). Structure of the biomedical literature. *Journal of the American Society for Information Science*, *27*(1), 25-45.

National Institutes of Health. (n.d.). *Clinical and translational science awards Progress Report 2009 - 2011*. Retrieved https://ncats.nih.gov/ctsa_2011/ch1.html

National Institutes of Health. (2017a). *ExPORTER Data Catalog*. Retrieved from https://exporter.nih.gov/ExPORTER_Catalog.aspx

National Institutes of Health. (2017b). *PAR-18-464: Clinical and Translational Science Award (U54)*. Retrieved from https://grants.nih.gov/grants/guide/pa-files/PAR-18-464.html

National Institutes of Health. (2017c). *What we Do*. Retrieved from https://www.nih.gov/about-nih/what-we-do

National Institutes of Health. (2018, November 18). *Research Portfolio Online Reporting Tools (RePORT)*. Retrieved from https://report.nih.gov.

National Science Foundation. (n.d.-a), *Engineering Research Centers*. Retrieved from http://erc-assoc.org

National Science Foundation. (n.d.-b), *Higher Education Research and Development Survey (HERD)*. Retrieved from https://www.nsf.gov/statistics/srvyherd/

National Science Foundation. (n.d.-c), *Survey of Graduate Students and Postdoctorates in Science and Engineering*. Retrieved from https://www.nsf.gov/statistics/srvygradpostdoc/

National Science Foundation (2018, February 2). *Classification of field of study*. Retrieved from https://www.nsf.gov/statistics/nsf13327/pdf/tabb1.pdf

Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of The National Academy of Sciences, 101*(suppl 1), 5200-5205.

Niosi, J. (2011). Complexity and path dependence in biotechnology innovation systems. *Industrial and Corporate Change, 20*(6), 1795-1826.

Northwestern University. (2019). *Northwestern University Clinical and Translational Sciences Institute*. Retrieved from https://www.nucats.northwestern.edu/funding/index.html

Obeid, J. S., Johnson, L. M., Stallings, S., & Eichmann, D. (2014). Research networking systems: the state of adoption at institutions aiming to augment translational research infrastructure. *Journal of translational medicine & epidemiology, 2*(2), 1026.

Ortega, L., & Antell, K. (2006). Tracking cross-disciplinary information use by author affiliation: Demonstration of a method. *College & Research Libraries, 67*(5), 446-462.

Owen-Smith, J., & Powell, W. W. (2003). The expanding role of university patenting in the life sciences: assessing the importance of experience and connectivity. *Research Policy*, *32*(9), 1695-1711.

Park, H. W., Yoon, J., & Leydesdorff, L. (2016). The normalization of co-authorship networks in the bibliometric evaluation: the government stimulation programs of China and Korea. *Scientometrics, 109*(2), 1017-1036.

PatentsView. (2018, November 27). *Data Download Tables*. Retrieved from http://www.patentsview.org/download

Porter, A., & Chubin, D. (1985). An indicator of cross-disciplinary research. *Scientometrics, 8*(3-4), 161-176.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 116-145.

Powell, W. W., White, D. R., Koput, K. W., & Owen-Smith, J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology, 110*(4), 1132-1205.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717-731.

Price, D. (1986). *Little science, big science-- and beyond*. New York, NY: Columbia University Press.

Qin, J., Lancaster, F. W., & Allen, B. (1997). Types and levels of collaboration in interdisciplinary research in the sciences. *JASIS, 48*(10), 893-916.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy, 41*(7), 1262-1282.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks, 29*(2), 173-191.

Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., ... & Goldberg, A. L. (1994). Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell, 78*(5), 761-771.

Rockey, S. (2012, January 6), *Working Through the Transitions of NCRR and NCATS*. Retrieved from https://nexus.od.nih.gov/all/2012/01/06/working-through-the-transitions-of-ncrr-and-ncats/

Roy, A., & Chaguturu, R. (2014). Current Trends in Collaborative Drug Discovery and Strategies to De-Risk Precompetitive Initiatives. In Chaguturu, R (Ed.), *Collaborative Innovation in Drug Discovery* (pp. 55-73). Hoboken, NJ: Wiley.

Rubio, D. M., Schoenbaum, E. E., Lee, L. S., Schteingart, D. E., Marantz, P. R., Anderson, K. E., ... & Esposito, K. (2010). Defining translational research: implications for training. *Academic Medicine, 85*(3), 470-475.

Sampat, B. N., & Pincus, H. A. (2015). Citations in life science patents to publicly funded research at academic medical centers. *Clinical and Translational Science, 8*(6), 759-763.

Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery, 11*(3), 191-200.

Schaefer, D. R. & Marcum, C. S. (2017, August 15). *Modeling Network Dynamics*. Retrieved from https://doi.org/10.31235/osf.io/6rm9q

Schneider, M., Kane, C. M., Rainwater, J., Guerrero, L., Tong, G., Desai, S. R., & Trochim, W. (2017). Feasibility of common bibliometrics in evaluating translational science. *Journal of Clinical and Translational Science, 1*(1), 45-52.

Schummer, J. (2004). Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics, 59*(3), 425-465.

Slagle, M. (2016, July 5). *New NIH-Funded center to study inefficiencies in clinical trials*. Retrieved from https://today.duke.edu/2016/07/trialinnovation

Statista (2019). *Research and development expenditure of total U.S. pharmaceutical industry from 1995 to 2017 (in billion U.S. dollars)*. Retrieved from https://www.statista.com/statistics/265085/research-and-development-expenditure-us-pharmaceutical-industry/

Steinmetz, K. L., & Spack, E. G. (2009). The basics of preclinical drug development for neurodegenerative disease indications. *BMC Neurology, 9*(1), S2.

Stevenson, D. K., Shaw, G. M., Wise, P. H., Norton, M. E., Druzin, M. L., Valantine, H. A., & McFarland, D. A. (2013). Transdisciplinary translational science and the case of preterm birth. *Journal of Perinatology, 33*(4), 251.

Stokes, T. D., & Hartley, J. A. (1989). Coauthorship, social structure and influence within specialties. *Social Studies of Science, 19*(1), 101-125.

Sung, H. Y., Wang, C. C., Huang, M. H., & Chen, D. Z. (2015). Measuring science-based science linkage and non-science-based linkage of patents through non-patent references. *Journal of Informetrics, 9*(3), 488-498.

Surkis, A., Hogle, J. A., DiazGranados, D., Hunt, J. D., Mazmanian, P. E., Connors, E., ... & Aphinyanaphongs, Y. (2016). Classifying publications from the clinical and translational science award program along the translational research spectrum: a machine learning approach. *Journal of Translational Medicine, 14*(1), 235.

Swan, J., Goussevskaia, A., Newell, S., Robertson, M., Bresnen, M., & Obembe, A. (2007). Modes of organizing biomedical innovation in the UK and US and the role of integrative and relational capabilities. *Research Policy, 36*(4), 529-547.

The President's Council of Advisors on Science and Technology. (2012). *Executive Summary Report to the President on Propelling Innovation in Drug Discovery, Development and Evaluation*. Retrieved from https://www.brookings.edu/wp-content/uploads/2012/11/PCAST-Executive-Summary.pdf

Translational Research Informatics Center. (2003). *Translational Research Informatics Center*. Retrieved from https://www.tri-kobe.org/english/index.html

U.S. Food and Drug Administration. (2018, January 4). *The Drug Development Process*. Retrieved from https://www.fda.gov/forpatients/approvals/drugs/

U.S. National Library of Medicine. (2017, June 3). *Translation Medical Research MeSH Descriptor Data 2017*. Retrieved from https://meshb.nlm.nih.gov/record/ui?name=Translational%20Medical%20Research

Van Alstyne, M., & Brynjolfsson, E. (1996). Could the Internet balkanize science? *Science, 274*(5292), 1479-1480.

Van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy,* 40(3), 463-472.

VanderWeele, T. J., & An, W. (2013). Social networks and causal inference. In S. L. Morgan (Ed.), *Handbook of causal analysis for social research (pp. 353-374).* New York, NY: Springer.

Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., ... & Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of informetrics, 5*(1), 14-26.

Walsh, J. P., & Maloney, N. G. (2007). Collaboration structure, communication media, and problems in scientific work teams. *Journal of Computer-Mediated Communication, 12*(2), 712-732.

Wang, X., Wang, Z., Huang, Y., Chen, Y., Zhang, Y., Ren, H., ... & Pang, J. (2017). Measuring interdisciplinarity of a research system: detecting distinction between publication categories and citation categories. *Scientometrics, 111*(3), 2023-2039.

Weber, G. M. (2013). Identifying translational science within the triangle of biomedicine. *Journal of Translational Medicine, 11*(1), 126.

Wild, H., Huwe, C., & Lessl, M. (2013). "Collaborative innovation"—regaining the edge in drug discovery. *Angewandte Chemie International Edition, 52*(10), 2684-2687.

Wilhelm, S., Tavares, A. J., Dai, Q., Ohta, S., Audet, J., Dvorak, H. F., & Chan, W. C. (2016). Analysis of nanoparticle delivery to tumours. *Nature Reviews Materials, 1*(5), 16014.

Williams, S. (2016). Building a translational medicine powerhouse in China, *Science*, 352(6293), 1596-1596.

Woolf, S. H. (2008). The meaning of translational research and why it matters. *JAMA*, 299(2), 211-213.

Working Group of the NCATS Advisory Council to the Director. (2014). *NCATS Advisory Council Working Group on the IOM Report: The CTSA Program at NIH*. Retrieved from https://ncats.nih.gov/files/CTSA-IOM-WG-report-5-2014.pdf

Wowk, K., McKinney, L., Muller-Karger, F., Moll, R., Avery, S., Escobar-Briones, E., ... & McLaughlin, R. (2017). Evolving academic culture to meet societal needs. *Palgrave Communications, 3*(1), 35.

Zerhouni, E. A. (2003). The NIH Roadmap. *Science, 302*(5642), 63-72.

Zerhouni, E. A. (2005). US biomedical research: basic, translational, and clinical sciences. *JAMA*, 294(11), 1352-1358.

Zerhouni, E. A. (2006). NIH in the post-doubling era: realities and strategies. *Science, 314*(5802), 1088-1090

Zettler, P. J., & Greely, H. T. (2014). The strange allure of state "right-to-try" laws. *JAMA Internal Medicine, 174*(12), 1885-1886.